GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 12:31:25 ; Search time 18725 Seconds
(without alignments)
11827.081 Million cell updates/sec

Title: US-09-828-068-1
Perfect score: 3896
Sequence: 1 cgcggctgtcggagcaaacg.....gtcaacaccggagaatttac 3896

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3892.8	99.9	3896	15	AF326768	AF326768 Oryza sat
	2	3882.4	99.7	4310	6	BD083699	BD083699 Novel gen
	3	3216.8	82.6	3626	15	AK103364	AK103364 Oryza sat
	4	2952.2	75.8	9455	6	BD083700	BD083700 Novel gen
	5	2952.2	75.8	110000	15	AP008207_071	Continuation (72 o
	6	2952.2	75.8	150594	15	AP001859	AP001859 Oryza sat
c	7	62.8	1.6	7218	6	I66494	I66494 Sequence 14
c	8	62	1.6	2000	6	AX655393	AX655393 Sequence
	9	53.2	1.4	2000	6	AX655393	AX655393 Sequence
c	10	50.8	1.3	179286	5	BX897685	BX897685 Zebrafish
	11	49.8	1.3	178247	5	BX248504	BX248504 Zebrafish
	12	48.8	1.3	110000	14	CT009752_4	Continuation (5 of
	13	48.8	1.3	110000	14	CT009752_5	Continuation (6 of
c	14	48.8	1.3	185541	5	BX005453	BX005453 Zebrafish
	15	48.6	1.2	228730	14	AC096088	AC096088 Rattus no
	16	48.4	1.2	124927	5	BX510987	BX510987 Zebrafish
	17	48.4	1.2	194784	5	BX890608	BX890608 Zebrafish
c	18	48.4	1.2	250029	2	AE014820	AE014820 Plasmodiu
	19	48.2	1.2	141041	14	AC026135	AC026135 Homo sapi
c	20	48.2	1.2	148130	8	AC092925	AC092925 Homo sapi
c	21	48.2	1.2	192029	14	AC113076	AC113076 Mus muscu
c	22	47.8	1.2	163843	5	BX510939	BX510939 Zebrafish
c	23	47.2	1.2	106172	8	AL591122	AL591122 Human DNA
	24	47.2	1.2	110000	15	AE016820_05	Continuation (6 of
	25	46.8	1.2	225453	14	AC097957	AC097957 Rattus no
	26	46.8	1.2	264572	14	AC097681	AC097681 Rattus no
c	27	46.6	1.2	128991	9	AC117775	AC117775 Mus muscu
c	28	46.6	1.2	187418	9	AL672278	AL672278 Mouse DNA
	29	46.4	1.2	142168	14	CR388191	CR388191 Danio rer
	30	46.2	1.2	34340	14	BX957360	BX957360 Danio rer
c	31	46.2	1.2	155304	8	AC055845	AC055845 Homo sapi
c	32	46.2	1.2	167446	14	AC137718	AC137718 Homo sapi
c	33	46.2	1.2	199203	14	AC026077	AC026077 Homo sapi
	34	45.6	1.2	79882	14	BX908758_5	Continuation (6 of
c	35	45.6	1.2	191146	5	BX530070	BX530070 Zebrafish
	36	45.4	1.2	274	6	AR248600	AR248600 Sequence
	37	45.4	1.2	145120	15	AP003434	AP003434 Oryza sat
c	38	45.2	1.2	1141	6	AR579680	AR579680 Sequence
c	39	45.2	1.2	1141	6	AX083744	AX083744 Sequence
	40	45.2	1.2	93100	8	AL162614	AL162614 Human DNA
c	41	45.2	1.2	94534	5	AL929250	AL929250 Zebrafish
	42	45.2	1.2	114980	14	AC087825	AC087825 Homo sapi
	43	45.2	1.2	158392	14	AC022048	AC022048 Homo sapi
	44	45	1.2	173233	8	AC116002	AC116002 Homo sapi
c	45	45	1.2	217726	9	AL928680	AL928680 Mouse DNA

ALIGNMENTS

RESULT 1
AF326768
LOCUS AF326768 3896 bp mRNA linear PLN 05-SEP-2001
DEFINITION Oryza sativa embryonic flower 1-like protein mRNA, complete cds.
ACCESSION AF326768
VERSION AF326768.1 GI:15430698
KEYWORDS .
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 3896)
AUTHORS Aubert,D., Chen,L., Moon,Y.H., Martin,D., Castle,L.A., Yang,C.H. and Sung,Z.R.
TITLE EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis
JOURNAL Plant Cell 13 (8), 1865-1875 (2001)
PUBMED 11487698
REFERENCE 2 (bases 1 to 3896)
AUTHORS Moon,Y.-H., Chen,L. and Sung,Z.R.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2000) Plant and Microbial Biol. Dept., University of California, Berkeley, 361 Koshland Hall, Berkeley, CA 94720, USA
FEATURES Location/Qualifiers
source 1. .3896
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
CDS 241. .3414
/note="EMF1"
/codon_start=1
/product="embryonic flower 1-like protein"
/protein_id="AAK98529.1"
/db_xref="GI:15430699"
/translation="MEIVAVDQEGARVVGTNCMLARGGTGAVAPVLELTATPRQDAAA
EAGVDEPAQHQCEHFSIRGYVALLQKKDPKFCSLSRIFHDQKKCDEHKASSSPFSVAK
FRRWDCSKCLDKLKTSDNGTAPRTLPAKQNGTSDGCSITFVRSTFVPASVGSQKVSPS
TQSSQGKNADRSTLPKSVQEGNDSKCNAPSGKNGAAEANTDSPMKDLQGPAQNYDVAA
NVSEDNTSVDVGALPEVPQITWHIEVNGADQPPSTPKLSEVVLKRNEDENGKTEETLV
AEQCNLTKDPNPMSGKERDQVAEQCNLTKDPKPVSGQKCEQICNEPCEEVVLKRSSKS

KRKTDKKLMKKQQHSKKRTAQADVSDAKLCRRKPKKVRLLSEIVNANQVEDSRSDEVH
RENAADPCEDDRSTIPVPMEVSMDIPVSNHTVGEDGLKSSKNKTKRKYSDVVDDGSSL
MNWLNGKKKRTGSVHHTVAHPAGNLSNKKVTPTASTQHDDENDTENGLDTNMHKTDVC
QHVSEISTQRCSSKGKTAGLSKGKTHSAASTKYGGESTRNGQNIHVLSAEDQCQMETE
NSVLSHSAKVSPAEHDIQIMSDLHEQSLPKKKKKQKLEVTREKQTMIDDIPMDIVELL
AKNQHERQLMTETDCSDINRIQSKTTADDDCVIVAAKDGSDYASSVFDTNSQQKSLAS
QSTQKELQGHLALTTQESPHPQNFQSTQEQQTHLRMEEMVTIAASSPLFSHHDDQYIA
EAPTEHWGRKDAKKLTWEQFKATTRNSPAATCGAQFRPGIQAVDLTSTHVMGSSSNYA
SRQPVIAPLDRYAERAVNQVHARNFPSTIATMEASKLCDRRNAGQVVLYPKESMPATH
LLRMMDPSTLASFPNYGTSSRNQMESQLHNSQYAHNQYKGSTSTSYGSNLNGKIPLTF
EDLSRHQLHDLHRPLRPHPRVGVLGSLLQKEIANWSENCGTQSGYKLGVSTGITSHQM
NRKEHFEALNSGMFSAKWNALQLGSVSSSAYFLSARNSIAQSWTRGKGKMVHPLDRFV
RQDICITNKNPADFTTISNDNEYMDYR"

ORIGIN

Query Match 99.9%; Score 3892.8; DB 15; Length 3896;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3894; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60

Qy   61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120

Qy  121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180

Qy  181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240

Qy  241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300

Qy  301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360

Qy  361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420

Qy  421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480

Qy  481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540

Qy  541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600

Qy  601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660

Qy  661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720

Qy  721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780

Qy  781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840

Qy  841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900

Qy  901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960

Qy  961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020

Qy 1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080

Qy 1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140

Qy 1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200

Qy 1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260

Qy 1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320

Qy 1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db 1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTGTAAATGCT 1380

Qy 1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440

Qy 1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500

Qy 1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560

Qy 1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620

Qy 1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680

Qy 1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740

Qy 1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800

Qy 1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860

Qy 1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
```

```
Qy       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980

Qy       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040

Qy       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100

Qy       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160

Qy       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220

Qy       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280

Qy       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340

Qy       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400

Qy       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460

Qy       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520

Qy       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580

Qy       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640

Qy       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700

Qy       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760

Qy       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820

Qy       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880

Qy       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940

Qy       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000

Qy       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060

Qy       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120

Qy       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180

Qy       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240

Qy       3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGTGCATATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300

Qy       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360

Qy       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3420

Qy       3421 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3480

Qy       3481 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3540

Qy       3541 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 3600

Qy       3601 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 3660

Qy       3661 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTC 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTC 3720

Qy       3721 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 3780

Qy       3781 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 3840

Qy       3841 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
```

RESULT 2
BD083699
LOCUS BD083699 4310 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel gene participating in response to brassinosteroid.
ACCESSION BD083699
VERSION BD083699.1 GI:22629309
KEYWORDS JP 2001327287-A/1.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

```
REFERENCE   1  (bases 1 to 4310)
  AUTHORS   Hirochika,H., Yamazaki,T. and Miyao,A.
  TITLE     Novel gene participating in response to brassinosteroid
  JOURNAL   Patent: JP 2001327287-A 1 27-NOV-2001;
            NATL INST OF AGROBIOLOGICAL RESOUCES, BIO ORIENTED TECHNOLOGY
            RESEARCH ADVA JAPAN FORAGE SEED ASSOCIATION
COMMENT     OS   Oryza sativa (rice)
            PN   JP 2001327287-A/1
            PD   27-NOV-2001
            PF   19-MAY-2000 JP 2000149106
            PI   HIROHIKO HIROCHIKA,TOKIO YAMAZAKI,AKIO MIYAO
            PC   C12N15/09,C12N15/00
            CC   Novel gene participating in response to brassinosteroid FH
            Key             Location/Qualifiers
            FT   CDS             (655). .(3828).
FEATURES             Location/Qualifiers
     source          1. .4310
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
ORIGIN

  Query Match           99.7%;  Score 3882.4;  DB 6;  Length 4310;
  Best Local Similarity 99.9%;  Pred. No. 0;
  Matches 3894;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 474

Qy         61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 534

Qy        121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 594

Qy        181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        595 TGAGTTGTGCCTCGGCTGTGCTAGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 654

Qy        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 714

Qy        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 774

Qy        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 834

Qy        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 894

Qy        481 ATTTTCCATGACCAGAAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 ATTTTCCATGACCAGAAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 954

Qy        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1014

Qy        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1074

Qy        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1134

Qy        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1194

Qy        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1254

Qy        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1314

Qy        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1374

Qy        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1434

Qy       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1494

Qy       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1554

Qy       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1614

Qy       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1674

Qy       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1734

Qy       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1735 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1794

Qy       1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1795 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1854

Qy       1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1914

Qy       1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1974

Qy       1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2034

Qy       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2094

Qy       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2154

Qy       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2214
```

```
Qy       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2274

Qy       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2334

Qy       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2394

Qy       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2454

Qy       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2514

Qy       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2574

Qy       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2634

Qy       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2694

Qy       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2695 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2754

Qy       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2755 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2814

Qy       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2815 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2874

Qy       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2875 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2934

Qy       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2935 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2994

Qy       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2995 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3054

Qy       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3114

Qy       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3174

Qy       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3234

Qy       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3235 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3294

Qy       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3295 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3354

Qy       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3355 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3414

Qy       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3415 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3474

Qy       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3534

Qy       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3594

Qy       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3595 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3654

Qy       3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3655 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3714

Qy       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3715 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3774

Qy       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3775 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3834

Qy       3421 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3835 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3894

Qy       3481 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3895 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3954

Qy       3541 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3955 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 4014

Qy       3601 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4015 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 4074

Qy       3661 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTC 3720
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       4075 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTG-AAAAAATGTCTC 4133

Qy       3721 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4134 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 4193

Qy       3781 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4194 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 4253

Qy       3841 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4254 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 4309
```

RESULT 3
AK103364
LOCUS AK103364 3626 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033126N23, full insert sequence.

ACCESSION AK103364
VERSION AK103364.1 GI:32988573
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
PUBMED 12869764
REFERENCE 2 (bases 1 to 3626)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.
FEATURES Location/Qualifiers
source 1. .3626
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033126N23"
ORIGIN

Query Match 82.6%; Score 3216.8; DB 15; Length 3626;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3218; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60
Db     407 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 466

Qy      61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120
Db     467 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 526

Qy     121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180
Db     527 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 586

Qy     181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240
Db     587 TGAGTTGTGCCTCGGCTGTGCTAGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 646

Qy     241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300
Db     647 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 706

Qy     301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360
Db     707 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 766

Qy     361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420
Db     767 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 826

Qy     421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480
Db     827 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 886

Qy     481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540
Db     887 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 946

Qy     541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600
Db     947 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1006

Qy     601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660
Db    1007 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1066

Qy     661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720
Db    1067 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1126
```

```
Qy    721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1127 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1186

Qy    781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1187 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1246

Qy    841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1247 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1306

Qy    901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1307 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1366

Qy    961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1367 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1426

Qy   1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1427 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1486

Qy   1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1546

Qy   1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1606

Qy   1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1666

Qy   1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1726

Qy   1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1786

Qy   1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1846

Qy   1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1906

Qy   1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1966

Qy   1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2026

Qy   1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2086

Qy   1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2146

Qy   1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2206
```

```
Qy   1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2266

Qy   1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2326

Qy   1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2386

Qy   1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2446

Qy   2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2447 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2506

Qy   2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2566

Qy   2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2626

Qy   2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2686

Qy   2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2687 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2746

Qy   2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2747 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2806

Qy   2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2807 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2866

Qy   2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2867 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2926

Qy   2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2927 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATATGGTGCTCAATTTAGACCTGGTATC 2986

Qy   2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2987 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3046

Qy   2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3047 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3106

Qy   2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3107 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3166

Qy   2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3167 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3226

Qy   2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3227 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3286

Qy   2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3287 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3346

Qy       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3347 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3406

Qy       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3407 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3466

Qy       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3467 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3526

Qy       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3527 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3586

Qy       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGC 3220
                ||||||||||||||||||||||||||||||||||||||||
Db       3587 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGC 3626


RESULT 4
BD083700
LOCUS       BD083700                9455 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Novel gene participating in response to brassinosteroid.
ACCESSION   BD083700
VERSION     BD083700.1  GI:22629310
KEYWORDS    JP 2001327287-A/2.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 9455)
  AUTHORS   Hirochika,H., Yamazaki,T. and Miyao,A.
  TITLE     Novel gene participating in response to brassinosteroid
  JOURNAL   Patent: JP 2001327287-A 2 27-NOV-2001;
            NATL INST OF AGROBIOLOGICAL RESOUCES, BIO ORIENTED TECHNOLOGY
            RESEARCH ADVANCEMENT INSTITUTION
COMMENT     OS   Oryza sativa (rice)
            PN   JP 2001327287-A/2
            PD   27-NOV-2001
            PF   19-MAY-2000 JP 2000149106
            PI   HIROHIKO HIROCHIKA,TOKIO YAMAZAKI,AKIO MIYAO
            PC   C12N15/09,C12N15/00
            CC   Novel gene participating in response to brassinosteroid FH
            Key             Location/Qualifiers
            FT   source          1. .9455
            FT                   /organism='Oryza sativa (rice)'.
FEATURES             Location/Qualifiers
     source          1. .9455
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
ORIGIN

  Query Match             75.8%;  Score 2952.2;  DB 6;  Length 9455;
  Best Local Similarity   87.4%;  Pred. No. 0;
  Matches 3487;  Conservative    0;  Mismatches   23;  Indels  481;  Gaps    4;

Qy        386 ACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGAGGGTATGTTGCTCTTCTTC 445
              | | | |   | ||    |||   |   ||  |  | ||||||||||||||||||||
Db       5447 ATGCATCTTTATAATGATAATCTTATTTTTCTTGTACCAGAGGGTATGTTGCTCTTCTTC 5506

Qy        446 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5507 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 5566

Qy        506 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       5567 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 5626

Qy        566 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5627 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 5686

Qy        626 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5687 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 5746

Qy        686 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5747 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 5806

Qy        746 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5807 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 5866

Qy        806 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA---------- 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5867 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGGTATGGTAG 5926

Qy        855 ---------------------------------------------------------- 854

Db       5927 ATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGAATAAACTAATGT 5986

Qy        855 ---------------AGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 899
                              ||||||||||||||||||||||||||||||||||||||||||||
Db       5987 TTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 6046

Qy        900 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6047 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 6106

Qy        960 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6107 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 6166

Qy       1020 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6167 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 6226

Qy       1080 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6227 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 6286

Qy       1140 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6287 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 6346

Qy       1200 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 1259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6347 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 6406

Qy       1260 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 1319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6407 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 6466

Qy       1320 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 1379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6467 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 6526

Qy       1380 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 1439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6527 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 6586

Qy       1440 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 1499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6587 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 6646

Qy       1500 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 1559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6647 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 6706
```

```
Qy      1560 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 1619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6707 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 6766

Qy      1620 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 1679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6767 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 6826

Qy      1680 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 1739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6827 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 6886

Qy      1740 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 1799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6887 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 6946

Qy      1800 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 1859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6947 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 7006

Qy      1860 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7007 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 7066

Qy      1920 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAA---------------- 1963
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db      7067 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTACGAATTTTGTG 7126

Qy      1964 ------------------------------------------------------------ 1963

Db      7127 AATCATGAGGAATTTTTGCTTTTTAAATTGACTGAATCAACATTTATCTGTATGAAGGAA 7186

Qy      1964 ------------------------------------------------------------ 1963

Db      7187 TAATATTGGTGCATAACAATGTTAAGAAATATGCATACAATGTTTATTTATATGCTTTCC 7246

Qy      1964 ------------------------------------------------------------ 1963

Db      7247 ACTGTTCTTCTTTACTTATGTTTTGATACTCTTTTTGTGTGTGCGTGCATGTGTGCATGT 7306

Qy      1964 ------------------------------------------------------------ 1963

Db      7307 GTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGCACGTGCGTGGCGCAATATTCTTTTTT 7366

Qy      1964 ----------------------------------------------------------AGGTTT 1969
                                                                       ||||||
Db      7367 AGACTCATATTATAGTGATTGTAATGGACTGACATTTTCCTCATTTCTCATCTCAGGTTT 7426

Qy      1970 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 2029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7427 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 7486

Qy      2030 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 2089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7487 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 7546

Qy      2090 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 2149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7547 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 7606

Qy      2150 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 2209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7607 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 7666

Qy      2210 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 2269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7667 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 7726

Qy      2270 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 2329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7727 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 7786
```

```
Qy      2330 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 2389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7787 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 7846

Qy      2390 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 2449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7847 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 7906

Qy      2450 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 2509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7907 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 7966

Qy      2510 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 2569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7967 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 8026

Qy      2570 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 2629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8027 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 8086

Qy      2630 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 2689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8087 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 8146

Qy      2690 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 2749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8147 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 8206

Qy      2750 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 2809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8207 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 8266

Qy      2810 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 2869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8267 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 8326

Qy      2870 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 2929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8327 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 8386

Qy      2930 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 2989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8387 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 8446

Qy      2990 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 3049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8447 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 8506

Qy      3050 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 3109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8507 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 8566

Qy      3110 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 3169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8567 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 8626

Qy      3170 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 3229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8627 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 8686

Qy      3230 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 3289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8687 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 8746

Qy      3290 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 3349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8747 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 8806

Qy      3350 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 3409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8807 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 8866

Qy      3410 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 3469
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8867 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 8926

Qy       3470 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTA-------- 3521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8927 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGTAAGT 8986

Qy       3522 ------------------------------------------------------------ 3521

Db       8987 TGAACTTTTTCTTCTTGCAAGTTTATCAGTTTAAGAAAAAAGAATGATTACTTATGTTAG 9046

Qy       3522 ----------------AGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 3565
                               ||||||||||||||||||||||||||||||||||||||||||||
Db       9047 CAAGGATGGTTCTTGCAGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 9106

Qy       3566 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 3625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9107 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 9166

Qy       3626 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 3685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9167 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 9226

Qy       3686 GGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 3745
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       9227 GGAGAAATGTGTAACCTTGTTG-AAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 9285

Qy       3746 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 3805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9286 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 9345

Qy       3806 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 3865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9346 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 9405

Qy       3866 TTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
              ||||||||||||||||||||||||||||||
Db       9406 TTACTCTTGAAGTCAACACCGGAGAATTTAC 9436


RESULT 5
AP008207_071
WPCOMMENT
  Sequence split into 433 fragments  LOCUS AP008207 Accession AP008207
     Fragment Name      Begin     End
     AP008207_000           1    110000
     AP008207_001      100001    210000
     AP008207_002      200001    310000
     AP008207_003      300001    410000
     AP008207_004      400001    510000
     AP008207_005      500001    610000
     AP008207_006      600001    710000
     AP008207_007      700001    810000
     AP008207_008      800001    910000
     AP008207_009      900001   1010000
     AP008207_010     1000001   1110000
     AP008207_011     1100001   1210000
     AP008207_012     1200001   1310000
     AP008207_013     1300001   1410000
     AP008207_014     1400001   1510000
     AP008207_015     1500001   1610000
     AP008207_016     1600001   1710000
     AP008207_017     1700001   1810000
     AP008207_018     1800001   1910000
     AP008207_019     1900001   2010000
     AP008207_020     2000001   2110000
     AP008207_021     2100001   2210000
     AP008207_022     2200001   2310000
     AP008207_023     2300001   2410000
     AP008207_024     2400001   2510000
     AP008207_025     2500001   2610000
     AP008207_026     2600001   2710000
     AP008207_027     2700001   2810000
     AP008207_028     2800001   2910000
     AP008207_029     2900001   3010000
     AP008207_030     3000001   3110000
     AP008207_031     3100001   3210000
     AP008207_032     3200001   3310000
     AP008207_033     3300001   3410000
     AP008207_034     3400001   3510000
     AP008207_035     3500001   3610000
     AP008207_036     3600001   3710000
     AP008207_037     3700001   3810000
     AP008207_038     3800001   3910000
     AP008207_039     3900001   4010000
     AP008207_040     4000001   4110000
     AP008207_041     4100001   4210000
     AP008207_042     4200001   4310000
     AP008207_043     4300001   4410000
     AP008207_044     4400001   4510000
     AP008207_045     4500001   4610000
     AP008207_046     4600001   4710000
     AP008207_047     4700001   4810000
     AP008207_048     4800001   4910000
     AP008207_049     4900001   5010000
     AP008207_050     5000001   5110000
     AP008207_051     5100001   5210000
     AP008207_052     5200001   5310000
     AP008207_053     5300001   5410000
     AP008207_054     5400001   5510000
     AP008207_055     5500001   5610000
     AP008207_056     5600001   5710000
     AP008207_057     5700001   5810000
     AP008207_058     5800001   5910000
     AP008207_059     5900001   6010000
     AP008207_060     6000001   6110000
     AP008207_061     6100001   6210000
     AP008207_062     6200001   6310000
     AP008207_063     6300001   6410000
     AP008207_064     6400001   6510000
     AP008207_065     6500001   6610000
     AP008207_066     6600001   6710000
     AP008207_067     6700001   6810000
     AP008207_068     6800001   6910000
     AP008207_069     6900001   7010000
     AP008207_070     7000001   7110000
     AP008207_071     7100001   7210000
     AP008207_072     7200001   7310000
     AP008207_073     7300001   7410000
     AP008207_074     7400001   7510000
     AP008207_075     7500001   7610000
     AP008207_076     7600001   7710000
     AP008207_077     7700001   7810000
     AP008207_078     7800001   7910000
     AP008207_079     7900001   8010000
     AP008207_080     8000001   8110000
     AP008207_081     8100001   8210000
     AP008207_082     8200001   8310000
     AP008207_083     8300001   8410000
     AP008207_084     8400001   8510000
     AP008207_085     8500001   8610000
     AP008207_086     8600001   8710000
     AP008207_087     8700001   8810000
     AP008207_088     8800001   8910000
     AP008207_089     8900001   9010000
     AP008207_090     9000001   9110000
     AP008207_091     9100001   9210000
     AP008207_092     9200001   9310000
     AP008207_093     9300001   9410000
     AP008207_094     9400001   9510000
     AP008207_095     9500001   9610000
     AP008207_096     9600001   9710000
     AP008207_097     9700001   9810000
     AP008207_098     9800001   9910000
     AP008207_099     9900001  10010000
     AP008207_100    10000001  10110000
```

AP008207_101	10100001	10210000
AP008207_102	10200001	10310000
AP008207_103	10300001	10410000
AP008207_104	10400001	10510000
AP008207_105	10500001	10610000
AP008207_106	10600001	10710000
AP008207_107	10700001	10810000
AP008207_108	10800001	10910000
AP008207_109	10900001	11010000
AP008207_110	11000001	11110000
AP008207_111	11100001	11210000
AP008207_112	11200001	11310000
AP008207_113	11300001	11410000
AP008207_114	11400001	11510000
AP008207_115	11500001	11610000
AP008207_116	11600001	11710000
AP008207_117	11700001	11810000
AP008207_118	11800001	11910000
AP008207_119	11900001	12010000
AP008207_120	12000001	12110000
AP008207_121	12100001	12210000
AP008207_122	12200001	12310000
AP008207_123	12300001	12410000
AP008207_124	12400001	12510000
AP008207_125	12500001	12610000
AP008207_126	12600001	12710000
AP008207_127	12700001	12810000
AP008207_128	12800001	12910000
AP008207_129	12900001	13010000
AP008207_130	13000001	13110000
AP008207_131	13100001	13210000
AP008207_132	13200001	13310000
AP008207_133	13300001	13410000
AP008207_134	13400001	13510000
AP008207_135	13500001	13610000
AP008207_136	13600001	13710000
AP008207_137	13700001	13810000
AP008207_138	13800001	13910000
AP008207_139	13900001	14010000
AP008207_140	14000001	14110000
AP008207_141	14100001	14210000
AP008207_142	14200001	14310000
AP008207_143	14300001	14410000
AP008207_144	14400001	14510000
AP008207_145	14500001	14610000
AP008207_146	14600001	14710000
AP008207_147	14700001	14810000
AP008207_148	14800001	14910000
AP008207_149	14900001	15010000
AP008207_150	15000001	15110000
AP008207_151	15100001	15210000
AP008207_152	15200001	15310000
AP008207_153	15300001	15410000
AP008207_154	15400001	15510000
AP008207_155	15500001	15610000
AP008207_156	15600001	15710000
AP008207_157	15700001	15810000
AP008207_158	15800001	15910000
AP008207_159	15900001	16010000
AP008207_160	16000001	16110000
AP008207_161	16100001	16210000
AP008207_162	16200001	16310000
AP008207_163	16300001	16410000
AP008207_164	16400001	16510000
AP008207_165	16500001	16610000
AP008207_166	16600001	16710000
AP008207_167	16700001	16810000
AP008207_168	16800001	16910000
AP008207_169	16900001	17010000
AP008207_170	17000001	17110000
AP008207_171	17100001	17210000
AP008207_172	17200001	17310000
AP008207_173	17300001	17410000
AP008207_174	17400001	17510000
AP008207_175	17500001	17610000
AP008207_176	17600001	17710000
AP008207_177	17700001	17810000
AP008207_178	17800001	17910000
AP008207_179	17900001	18010000
AP008207_180	18000001	18110000
AP008207_181	18100001	18210000
AP008207_182	18200001	18310000
AP008207_183	18300001	18410000
AP008207_184	18400001	18510000
AP008207_185	18500001	18610000
AP008207_186	18600001	18710000
AP008207_187	18700001	18810000
AP008207_188	18800001	18910000
AP008207_189	18900001	19010000
AP008207_190	19000001	19110000
AP008207_191	19100001	19210000
AP008207_192	19200001	19310000
AP008207_193	19300001	19410000
AP008207_194	19400001	19510000
AP008207_195	19500001	19610000
AP008207_196	19600001	19710000

Query Match 75.8%; Score 2952.2; DB 15; Length 110000;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 3487; Conservative 0; Mismatches 23; Indels 481; Gaps 4;

```
Qy      386 ACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGAGGGTATGTTGCTCTTCTTC 445
            | | | |   | ||    |||   |   ||  |  |  ||||||||||||||||||||||
Db    59396 ATGCATCTTTATAATGATAATCTTATTTTTCTTGTACCAGAGGGTATGTTGCTCTTCTTC 59455

Qy      446 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 505
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59456 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 59515

Qy      506 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 565
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59516 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 59575

Qy      566 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 625
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59576 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 59635

Qy      626 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 685
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59636 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 59695

Qy      686 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 745
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59696 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 59755

Qy      746 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 805
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59756 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 59815

Qy      806 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA---------- 854
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59816 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGGTATGGTAG 59875

Qy      855 ------------------------------------------------------------ 854

Db    59876 ATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGAATAAACTAATGT 59935

Qy      855 ---------------AGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 899
                            ||||||||||||||||||||||||||||||||||||||||||||||
Db    59936 TTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 59995

Qy      900 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 959
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    59996 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 60055

Qy      960 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 1019
```

```
Db      60056 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 60115
Qy       1020 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 1079
Db      60116 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 60175
Qy       1080 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 1139
Db      60176 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 60235
Qy       1140 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 1199
Db      60236 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 60295
Qy       1200 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 1259
Db      60296 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 60355
Qy       1260 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 1319
Db      60356 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 60415
Qy       1320 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 1379
Db      60416 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 60475
Qy       1380 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 1439
Db      60476 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 60535
Qy       1440 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 1499
Db      60536 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 60595
Qy       1500 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 1559
Db      60596 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 60655
Qy       1560 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 1619
Db      60656 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 60715
Qy       1620 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 1679
Db      60716 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 60775
Qy       1680 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 1739
Db      60776 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 60835
Qy       1740 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 1799
Db      60836 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 60895
Qy       1800 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 1859
Db      60896 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 60955
Qy       1860 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 1919
Db      60956 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 61015
Qy       1920 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAA---------------- 1963
Db      61016 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTACGAATTTTGTG 61075
Qy       1964 ------------------------------------------------------------ 1963
Db      61076 AATCATGAGGAATTTTTGCTTTTTAAATTGACTGAATCAACATTTATCTGTATGAAGGAA 61135
Qy       1964 ------------------------------------------------------------ 1963
```

```
Db      61136 TAATATTGGTGCATAACAATGTTAAGAAATATGCATACAATGTTTATTTATATGCTTTCC 61195
Qy       1964 ------------------------------------------------------------ 1963
Db      61196 ACTGTTCTTCTTTACTTATGTTTTGATACTCTTTTTGTGTGTGCGTGCATGTGTGCATGT 61255
Qy       1964 ------------------------------------------------------------ 1963
Db      61256 GTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGCACGTGCGTGGCGCAATATTCTTTTTT 61315
Qy       1964 ----------------------------------------------------------AGGTTT 1969
Db      61316 AGACTCATATTATAGTGATTGTAATGGACTGACATTTTCCTCATTTCTCATCTCAGGTTT 61375
Qy       1970 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 2029
Db      61376 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 61435
Qy       2030 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 2089
Db      61436 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 61495
Qy       2090 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 2149
Db      61496 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 61555
Qy       2150 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 2209
Db      61556 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 61615
Qy       2210 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 2269
Db      61616 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 61675
Qy       2270 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 2329
Db      61676 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 61735
Qy       2330 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 2389
Db      61736 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 61795
Qy       2390 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 2449
Db      61796 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 61855
Qy       2450 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 2509
Db      61856 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 61915
Qy       2510 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 2569
Db      61916 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 61975
Qy       2570 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 2629
Db      61976 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 62035
Qy       2630 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 2689
Db      62036 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 62095
Qy       2690 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 2749
Db      62096 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 62155
Qy       2750 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 2809
Db      62156 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 62215
Qy       2810 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 2869
Db      62216 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 62275
```

```
Qy       2870 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 2929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62276 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 62335

Qy       2930 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 2989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62336 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 62395

Qy       2990 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 3049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62396 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 62455

Qy       3050 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 3109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62456 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 62515

Qy       3110 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 3169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62516 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 62575

Qy       3170 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 3229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62576 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 62635

Qy       3230 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 3289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62636 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 62695

Qy       3290 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 3349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62696 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 62755

Qy       3350 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 3409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62756 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 62815

Qy       3410 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 3469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62816 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 62875

Qy       3470 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTA-------- 3521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62876 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGTAAGT 62935

Qy       3522 ---------------------------------------------------------- 3521

Db      62936 TGAACTTTTTCTTCTTGCAAGTTTATCAGTTTAAGAAAAAAGAATGATTACTTATGTTAG 62995

Qy       3522 ----------------AGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 3565
                               ||||||||||||||||||||||||||||||||||||||||||||
Db      62996 CAAGGATGGTTCTTGCAGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 63055

Qy       3566 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 3625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63056 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 63115

Qy       3626 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 3685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63116 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 63175

Qy       3686 GGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 3745
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      63176 GGAGAAATGTGTAACCTTGTTG-AAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 63234

Qy       3746 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 3805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63235 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 63294

Qy       3806 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 3865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63295 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 63354

Qy       3866 TTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
              ||||||||||||||||||||||||||||||
Db      63355 TTACTCTTGAAGTCAACACCGGAGAATTTAC 63385
```

RESULT 6
AP001859

```
LOCUS       AP001859               150594 bp    DNA     linear   PLN 19-JAN-2005
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
            PAC clone:P0485D09.
ACCESSION   AP001859 BA000010
VERSION     AP001859.1  GI:7630233
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
            Katayose,Y., Wu,J., Niimura,Y., Cheng,Z., Nagamura,Y.,
            Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
            Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
            Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
            Idonuma,A., Iijima,M., Ikeda,M., Ikeno,M., Ito,S., Ito,T., Ito,Y.,
            Ito,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katagiri,S.,
            Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
            Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
            Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
            Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
            Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
            Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
            Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
            Yano,M., Jiang,J. and Gojobori,T.
  TITLE     The genome sequence and structure of rice chromosome 1
  JOURNAL   Nature 420 (6913), 312-316 (2002)
   PUBMED   12447438
REFERENCE   2  (bases 1 to 150594)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki, 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'. A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene
            predicted by a single gene prediction program is also classified as
            a probable 'hypothetical' protein and is included as a
            miscellaneous feature of the sequence.
```

```
          The orientation of the sequence is from T7 to SP6 of the PAC clone.
          This sequence of P0485D09 clone has an overlap with P0452F10(DDBJ:
          AP003434) clone at 5' end and with P0431F01(DDBJ: AP001550) clone
          at 3' end. Detailed information on overlap and assembly quality
          together with annotation of this entry is available at
          http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1. .150594
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="1"
                     /clone="P0485D09"
     gene            complement(4801. .5133)
                     /gene="P0485D09.1"
     mRNA            complement(<4801. .>5133)
                     /gene="P0485D09.1"
                     /note="start and end point are not identified"
     CDS             complement(4801. .5133)
                     /gene="P0485D09.1"
                     /note="predicted by FGENESH etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD81318.1"
                     /db_xref="GI:56783881"
                     /translation="MTELRSGLGVAGTCAMAHGASCLITLSPQLKDYEDTFSLTGKIQ
                     NFRELLYYSDGLNSEEKRGSRWLGTAANDVSLSDFSPAATTNSPLIMLWSLANLVRLG
                     LVAPWISK"
     gene            join(5701. .5883,6903. .6908)
                     /gene="P0485D09.2"
     misc_feature    join(5701. .5883,6903. .6908)
                     /gene="P0485D09.2"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            join(7919. .7964,8092. .8280,8453. .8473,8558. .8610)
                     /gene="P0485D09.3"
     misc_feature    join(7919. .7964,8092. .8280,8453. .8473,8558. .8610)
                     /gene="P0485D09.3"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            complement(9235. .9582)
                     /gene="P0485D09.4"
     misc_feature    complement(9235. .9582)
                     /gene="P0485D09.4"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            10769. .17070
                     /gene="P0485D09.5"
     mRNA            join(<10769. .10884,10933. .10987,11596. .11653,12285. .12379,
                     12471. .12574,12958. .13086,13493. .15281,15497. .15702,
                     16513. .16597,16831. .>17070)
                     /gene="P0485D09.5"
                     /note="start and end point are not identified"
     CDS             join(10769. .10884,10933. .10987,11596. .11653,12285. .12379,
                     12471. .12574,12958. .13086,13493. .15281,15497. .15702,
                     16513. .16597,16831. .17070)
                     /gene="P0485D09.5"
                     /codon_start=1
                     /product="putative Myb-like protein"
                     /protein_id="BAD81319.1"
                     /db_xref="GI:56783882"
                     /translation="MESLLDLRRLLKKEKSAMNQKGDDLSMGKRLKALLTIVMRTTGP
                     TRRSTKGNWTPEEDAILSRAVQTYNGKNWKKIAECFPDRTDVQCLHRWQKVLNPELVK
                     GPWSKEEDEIIVQMVNKLGPKKWSTIAQALPGRIGKQCRERWYNHLNPGINKEAWTQE
                     EEITLIHAHRMYGNKWAELTKFLPGRTDNSIKNHWNSSVKKKVNSYMSSGLLTQVSCL
                     PLNEYSANCNSSPAMTQQNSEDSGCFAVREVENSSGCSQSSLAKVSCSQVHDTTVPLG
                     CDLQVNANFDKNEAHDSQSSMGPQACYTSAEAVASALPAVHCHVSSSNLDPDQHLQED
                     FAQGLNLDMTIDEMPTVPSFADNQTVCSIENHERSLEPYDVAMEVPLSMLSSDSGAEQ
                     KLHFMSEADFNSPNCLKSELWQDISLQGLLSGPDAVEADSISRSNHQSDVYSSEADTH
                     FLAPPYMPQTSNSSSVMGLADDQSPQMSVPPSLICSNAMTDDAPFDNRPGRKEMPLSQ
                     AEVVTQSSSSSGDAEMFANPGCSNDRHVPSSTMESIPECGDQQVTNAEEPEASLEKEP
                     SLTQSVTAPDEQDKGALFYEPPRFPSLDVPFVSCDLVTSGDLQEFSPLGIRQLMHSTM
                     NVCTPMRLWGSPTHDESTGVLLKSAAKSFICTPSILKKRHRDLLSPIPDKRIEKKYGT
                     EKDRGVSDTSSTGIQTSCINATKDDALITTVLRIERSASSKSLEKKLVFSDENKENLG
                     YTTEQTKDGQSAGNDEHMDEQTTGERSSATNVATNDDLSGNLQPAGILIEHSGDDPIS
                     PDYGKNTMNQKLNTNVKSLSVCKEGVCAKSKPTELIVEKSSPCINVDYEYVNILADTP
                     GIKRGLESPSAWKSPWFVDMHFQGSYFTSPADSYDALGLMKQINVQTAAALVEAREVL
                     ASGGQCDNISSDKENTGNPDAKKEPGTTKLQTKIMVRGSELPLR"
     gene            complement(join(19054. .19378,19502. .19573,19905. .19998,
                     20072. .20220,20774. .21113,21407. .21738,21975. .22540))
                     /gene="P0485D09.6-1"
     mRNA            complement(join(19054. .19378,19502. .19573,19905. .19998,
                     20072. .20220,20774. .21113,21407. .21738,21975. .22540))
                     /gene="P0485D09.6-1"
                     /note="supported by full-length cDNA(s): AK070838"
     CDS             complement(join(19334. .19378,19502. .19573,19905. .19998,
                     20072. .20220,20774. .21113,21407. .21738,21975. .22508))
                     /gene="P0485D09.6-1"
                     /note="contains EST(s): AU173357(R2418)
                     contains full-length cDNA(s): AK070838"
                     /codon_start=1
                     /product="putative eukaryotic translation initiation
                     factor 3 subunit 6 interacting protein"
                     /protein_id="BAD81320.1"
                     /db_xref="GI:56783883"
                     /translation="MAAAAAEGGGGGGGAVYDPSYVPDSVKTFVAHMYRHVRDKNVYE
                     IHQMYEGGFQRLSDRLFRDAPWPAAEAVSPYCDGDHVFLLLYRELWYRHAYARASSSS
                     SSSAPLTAGQRAESWANYCDLFSVVLHGVVNMQLPNQWLWDMVDEFVYQFQSFCQYRA
                     KLKNKSDDELHQLKQFDKAWNVYGVLNYLQALVEKSMIAQILEREKEGLEQFTATDGY
                     DYQGGSNVLKMLGYYSMIGLLRIHCLLGDYRTGLKCLAPIDLNQQGVYTIVIGSHISA
                     IYHYGFANLMMRRYAEAIREFNKILLYILKYKQYHQKSPQYDQILKKNEQMYAFLAVC
                     LSLCPQHNLIDENVSTQLKEKYNDKMTKMQRFDEETYAAYDELFSYACPKFITPSPPA
                     LDQPLTNYNQDAYRLQLKLFLYEVKQQQLLSGIRSYLKLYSTITIAKLAQYMEVDEAT
                     LRSILMTYKHKMHAVDNNGKIVSSADFDFYIKEDVIHVMESKPIKRHGDYFLRQILKF
                     EEMIGELEKVQFD"
     gene            complement(join(21153. .21738,21975. .22547))
                     /gene="P0485D09.6-2"
     mRNA            complement(join(21153. .21738,21975. .22547))
                     /gene="P0485D09.6-2"
                     /note="supported by full-length cDNA(s): AK061285"
     CDS             complement(join(21400. .21738,21975. .22508))
                     /gene="P0485D09.6-2"
                     /note="contains full-length cDNA(s): AK061285"
                     /codon_start=1
                     /product="putative eukaryotic translation initiation
                     factor 3 subunit 6 interacting protein"
                     /protein_id="BAD81321.1"
                     /db_xref="GI:56783884"
                     /translation="MAAAAAEGGGGGGGAVYDPSYVPDSVKTFVAHMYRHVRDKNVYE
                     IHQMYEGGFQRLSDRLFRDAPWPAAEAVSPYCDGDHVFLLLYRELWYRHAYARASSSS
                     SSSAPLTAGQRAESWANYCDLFSVVLHGVVNMQLPNQWLWDMVDEFVYQFQSFCQYRA
                     KLKNKSDDELHQLKQFDKAWNVYGVLNYLQALVEKSMIAQILEREKEGLEQFTATDGY
                     DYQGGSNVLKMLGYYSMIGLLRIHCLLGDYRTGLKCLAPIDLNQQGVYTIVIGSHISA
                     IYHYGFANLMMRRY"
     gene            22882. .30975
                     /gene="P0485D09.7"
     mRNA            join(22882. .23145,24592. .24670,25118. .25216,25302. .25415,
                     25608. .25750,26364. .26445,27073. .27158,27527. .27760,
                     29332. .29810,29957. .30975)
                     /gene="P0485D09.7"

  Query Match           75.8%;  Score 2952.2;  DB 15;  Length 150594;
  Best Local Similarity  87.4%;  Pred. No. 0;
  Matches 3487;  Conservative    0;  Mismatches   23;  Indels  481;  Gaps    4;

Qy        386 ACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGAGGGTATGTTGCTCTTCTTC 445
              | | | |  | ||    |||  |  ||  | |  ||||||||||||||||||||||
Db      54044 ATGCATCTTTATAATGATAATCTTATTTTTCTTGTACCAGAGGGTATGTTGCTCTTCTTC 54103

Qy        446 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 505
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54104 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 54163
```

```
Qy       506 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 565
Db     54164 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 54223

Qy       566 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 625
Db     54224 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 54283

Qy       626 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 685
Db     54284 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 54343

Qy       686 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 745
Db     54344 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 54403

Qy       746 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 805
Db     54404 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 54463

Qy       806 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA---------- 854
Db     54464 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGGTATGGTAG 54523

Qy       855 ---------------------------------------------------------- 854
Db     54524 ATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGAATAAACTAATGT 54583

Qy       855 ----------AGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 899
Db     54584 TTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 54643

Qy       900 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 959
Db     54644 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 54703

Qy       960 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 1019
Db     54704 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 54763

Qy      1020 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 1079
Db     54764 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 54823

Qy      1080 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 1139
Db     54824 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 54883

Qy      1140 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 1199
Db     54884 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 54943

Qy      1200 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 1259
Db     54944 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 55003

Qy      1260 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 1319
Db     55004 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 55063

Qy      1320 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 1379
Db     55064 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 55123

Qy      1380 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 1439
Db     55124 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 55183

Qy      1440 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 1499
Db     55184 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 55243

Qy      1500 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 1559
Db     55244 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 55303

Qy      1560 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 1619
Db     55304 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 55363

Qy      1620 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 1679
Db     55364 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 55423

Qy      1680 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 1739
Db     55424 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 55483

Qy      1740 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 1799
Db     55484 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 55543

Qy      1800 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 1859
Db     55544 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 55603

Qy      1860 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 1919
Db     55604 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 55663

Qy      1920 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAA---------------- 1963
Db     55664 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTACGAATTTTGTG 55723

Qy      1964 ---------------------------------------------------------- 1963
Db     55724 AATCATGAGGAATTTTTGCTTTTTAAATTGACTGAATCAACATTTATCTGTATGAAGGAA 55783

Qy      1964 ---------------------------------------------------------- 1963
Db     55784 TAATATTGGTGCATAACAATGTTAAGAAATATGCATACAATGTTTATTTATATGCTTTCC 55843

Qy      1964 ---------------------------------------------------------- 1963
Db     55844 ACTGTTCTTCTTTACTTATGTTTTGATACTCTTTTTGTGTGTGCGTGCATGTGTGCATGT 55903

Qy      1964 ---------------------------------------------------------- 1963
Db     55904 GTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGCACGTGCGTGGCGCAATATTCTTTTTT 55963

Qy      1964 ----------------------------------------------------------AGGTTT 1969
Db     55964 AGACTCATATTATAGTGATTGTAATGGACTGACATTTTCCTCATTTCTCATCTCAGGTTT 56023

Qy      1970 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 2029
Db     56024 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 56083

Qy      2030 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 2089
Db     56084 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 56143

Qy      2090 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 2149
Db     56144 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 56203

Qy      2150 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 2209
Db     56204 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 56263

Qy      2210 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 2269
Db     56264 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 56323

Qy      2270 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 2329
```

```
Db      56324 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 56383

Qy       2330 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 2389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56384 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 56443

Qy       2390 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 2449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56444 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 56503

Qy       2450 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 2509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56504 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 56563

Qy       2510 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 2569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56564 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 56623

Qy       2570 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 2629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56624 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 56683

Qy       2630 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 2689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56684 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 56743

Qy       2690 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 2749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56744 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 56803

Qy       2750 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 2809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56804 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 56863

Qy       2810 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 2869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56864 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 56923

Qy       2870 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 2929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56924 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 56983

Qy       2930 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 2989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56984 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 57043

Qy       2990 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 3049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57044 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 57103

Qy       3050 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 3109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57104 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 57163

Qy       3110 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 3169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57164 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 57223

Qy       3170 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 3229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57224 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 57283

Qy       3230 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 3289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57284 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 57343

Qy       3290 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 3349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57344 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 57403

Qy       3350 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 3409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      57404 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 57463

Qy       3410 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 3469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57464 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 57523

Qy       3470 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTA-------- 3521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57524 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGTAAGT 57583

Qy       3522 ------------------------------------------------------------ 3521

Db      57584 TGAACTTTTTCTTCTTGCAAGTTTATCAGTTTAAGAAAAAAGAATGATTACTTATGTTAG 57643

Qy       3522 ----------------AGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 3565
                               ||||||||||||||||||||||||||||||||||||||||||||
Db      57644 CAAGGATGGTTCTTGCAGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 57703

Qy       3566 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 3625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57704 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 57763

Qy       3626 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 3685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57764 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 57823

Qy       3686 GGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 3745
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      57824 GGAGAAATGTGTAACCTTGTTG-AAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 57882

Qy       3746 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 3805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57883 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 57942

Qy       3806 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 3865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57943 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 58002

Qy       3866 TTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
             ||||||||||||||||||||||||||||||
Db      58003 TTACTCTTGAAGTCAACACCGGAGAATTTAC 58033
```

```
RESULT 7
I66494/c
LOCUS       I66494                 7218 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   I66494
VERSION     I66494.1  GI:2724471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 7218)
  AUTHORS   Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
  TITLE     Recombinant fowlpox virus
  JOURNAL   Patent: US 5670367-A 14 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .7218
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             1.6%;  Score 62.8;  DB 6;  Length 7218;
  Best Local Similarity   7.2%;  Pred. No. 0.00014;
  Matches   31;  Conservative  227;  Mismatches  174;  Indels    0;  Gaps    0;

Qy       1182 ATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATC 1241
             ||||  |||   ||| ||| | | || |    |||| |||    | | |
Db       1497 ATGTAGGCATCACTGTAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAAGAATTTGG 1438

Qy       1242 TAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGC 1301
             || :::::::: ::: ::::::  :: ::::::: :: :: : :: :::::: : :  :
```

```
Db       1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378

Qy       1302 CCAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCT 1361
Db       1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318

Qy       1362 ATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGA 1421
Db       1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258

Qy       1422 AAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAG 1481
Db       1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198

Qy       1482 CATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAA 1541
Db       1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138

Qy       1542 CAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCT 1601
Db       1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078

Qy       1602 GAATGGAAAAAA 1613
Db       1077 RRRRRRRRRRA 1066
```

RESULT 8
AX655393/c
LOCUS AX655393 2000 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 5263 from Patent WO03000898.
ACCESSION AX655393
VERSION AX655393.1 GI:29158207
KEYWORDS .
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M., Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
source 1. .2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN

Query Match 1.6%; Score 62; DB 6; Length 2000;
Best Local Similarity 10.5%; Pred. No. 0.00021;
Matches 82; Conservative 360; Mismatches 325; Indels 11; Gaps 3;

```
Qy       1904 TCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAA 1963
Db        939 WCWYCCMWKWWKMTSCMWWKYWRTWSCWYTMWWGAMRYAYYAMRRRRWTYKWSWRRMYWT 880

Qy       1964 AGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTAC 2023
Db        879 MTKWAWTWMTCMCMAKWYMATGWATWMWWRYTMYTYCYAMTCAKCKYKMAMTKWWTTWAC 820

Qy       2024 CCAAGAAGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATG 2083
Db        819 AWRATSWRWRAMAG----MRWKRYKMKRAYWWRWWRCWKAGWARWMKSRYRWKWKKYATR 764

Qy       2084 ACATCCCCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGA 2143
Db        763 YYWKMWAMTWWWSWRRWKSYRMWSGMGRMRWSAWRYCSRMKCAKTKYASSARWTKRAKRS 704

Qy       2144 CTGAGACTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATT 2203
Db         703 YRYRRRWYWKRKGWTYRYRYWRSCRMTRARMSKRRKWAGASMKSCWMYWRGARSMWYSKY 644

Qy       2204 GTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATT 2263
Db         643 SCSAKCCKKTRYMTSSYMSTGMYGMYSSYKSMSWTSKMSYMGKMTCTMYTSMKGSTRRSK 584

Qy       2264 C-----CCAACAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTT 2318
Db         583 MGRWSGMSRMYMRWWKKMRKRKYMRYMKWKCTWRRCMCYRWGYTMYTTSRSRMMYTGRYK 524

Qy       2319 GGCATTGACCACACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCA 2378
Db         523 ARYTSKRRYMWYKYRKYCWYYYYGMYMKCSYMMRYGYCKACKKCCYAM--CWKAAYSGMM 466

Qy       2379 GACACATTTGCGGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACA 2438
Db         465 MYWYRKYSKWMRMSTKYMWSMWYKKCRSMKYGAKGCYGCKMWTYCSYGYMKWYTYMGSYK 406

Qy       2439 TCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAA 2498
Db         405 YSRCYKYMRMYMYKGWMYMMYYSAYSSMMTWYYYYAKYWKYWYKRRGTMSWYGKSYKKKY 346

Qy       2499 GAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGG 2558
Db         345 CTWWCYMKCMRCYRWRKMMRKKTKYSKRCYCWRYATCYWCCCYRKRGWYSRRSMMRTAGK 286

Qy       2559 TGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATC 2618
Db         285 WKMRSWSRWCRSYSWYKMYKKMWKKSYYMSYGWARSSGTWSRSAAKRTYKGYSTSRRAKM 226

Qy       2619 TTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAA 2676
Db         225 MRACRMYSACRRYSRTSYYCGCSYCGSSKWKYMSKSCSMRMTCSSWCSCCYTCYYGAM 168

RESULT 9
AX655393
LOCUS AX655393 2000 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 5263 from Patent WO03000898.
ACCESSION AX655393
VERSION AX655393.1 GI:29158207
KEYWORDS .
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M., Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
source 1. .2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN

Query Match 1.4%; Score 53.2; DB 6; Length 2000;
Best Local Similarity 10.3%; Pred. No. 0.04;
Matches 95; Conservative 401; Mismatches 414; Indels 14; Gaps 2;

```
Qy   198 GTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGATATGGAGATTGTTGCAGT 257
Db   7 GWKSRCSSWRGRRRMYMAGMMSCARMGSSRMSRKMGSMSKYRKCSSCGKCKMTTRRKSKW 66

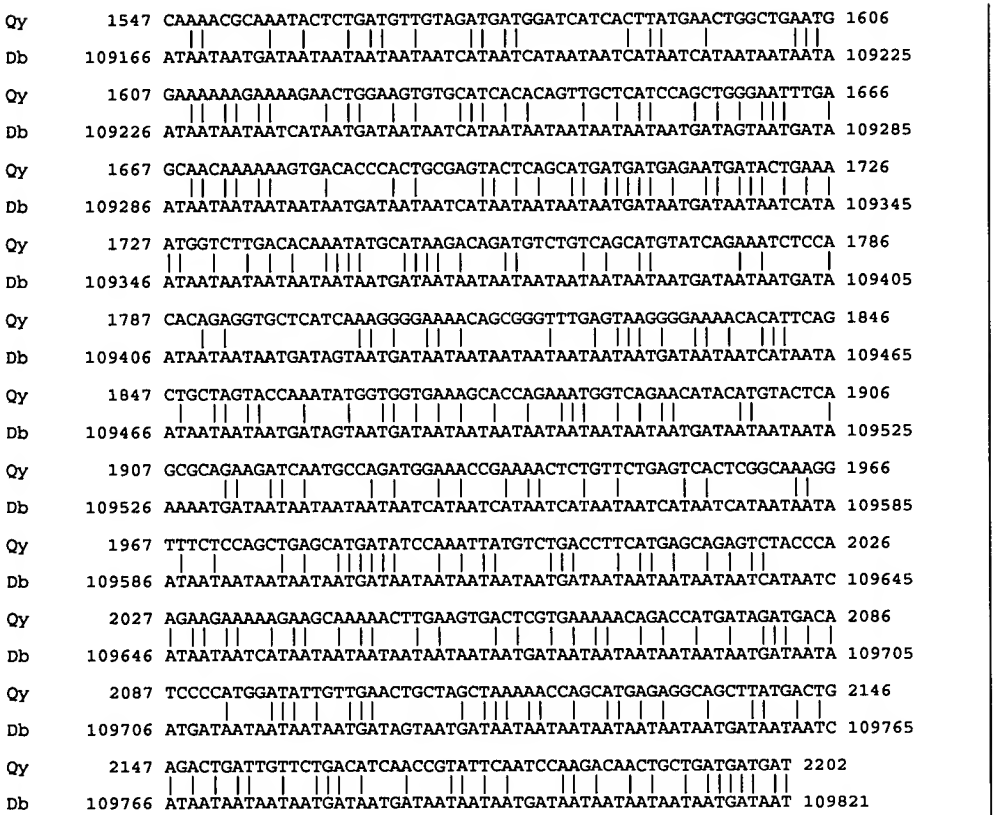
Qy   258 AGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTTGCTCGTGGTGGAACTGG 317
Db   67 YSASSASGRTGSKWSSGSYSGKGMKKRYKRSKRWRGRGRRGMRRSRMRWMGRYRRCARSG 126
```
```

```
Qy        318 TGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAGGATGCAGCCGCTGAAGC 377
Db        127 RMAGGSGRMMGGKSRMSYWMWCYARGCGSCKRKKSKGGSWGKTCRRGARGGSGWSSGAKY 186

Qy        378 TGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGAGGGTATGTTGC 437
Db        187 KSGSMSKRMWMSSCGRSGCGRRSAYSRYYGTSRKYGTYKKMTYYSASRCMRAYMTTSYSW 246

Qy        438 TCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAA 497
Db        247 ACSSYTWCRSKRRSMMWKMMRKMRWSRSYGWYSWSYKMWMCTAYKKSYYSRWCYMYRGGG 306

Qy        498 AAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATG 557
Db        307 WRGATRYWGRGYMSRMAMM----------YKKMYWYRGYKGMKRGWWAGRMMMRSMC 353

Qy        558 GGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAAC 617
Db        354 RWSKACYYMRWRMWRMTRRRRWAKKSSRTSRRKKRKWCMRKRKYKRMRGYSRMRSCKRAR 413

Qy        618 TCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCAC 677
Db        414 WMKRCRSGRAWKMGCRGCMTCRMKSYGMMRWKSWKRMASKYKWMSRMYRWRKKKCSRTTM 473

Qy        678 TTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGG 737
Db        474 WGKTRGGMMGTMGRCRYKKRSGMKRKCRRRRWGRMYRMRWKRYYMSARYTMRYCARKKYS 533

Qy        738 GAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATG 797
Db        534 YSAARKARCWYRGKGYYWAGMWMKRYKRMYMYKMMWWYKRKYSKCSWYCKMSYYASCMKS 593

Qy        798 CAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGA 857
Db        594 ARKAGAKMCKRSKMSAWSKSMRSSRKCRKCASKRSSAKRYAMMGGMTSGSRMSRWKSYTC 653

Qy        858 TTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTC 917
Db        654 YWRKWGSMKSTCTWMYYMSKYTYAKYGSYWRYRYRAWCMYMWRWYYYRYRSYMTYMAWYT 713

Qy        918 TGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGG 977
Db        714 SSTRMAMTGMKYSGRYWTSWYKYCKCSWKYRSMWYYWSWWWAKTWKMWRRYATRMMWMWY 773

Qy        978 TGCAGATCAACCTCCATCCACTCCAAAACTTTCTG-AAGTGGTCCTCAAAAGAAATGAAG 1036
Db        774 RYSMKWYTWCTMWGYWWYWWRTYMKMRYMWYKCTKTYWYWSATYWTGTWAAWWMAKTKMR 833

Qy       1037 ATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATC 1096
Db        834 MGMTGAKTRGRARKARYWWKWATWCATKRWMTKGKGAKWAWTWMAKAWRKYYWSWMRAWY 893

Qy       1097 CTAACCCAATGTCTGGAAAGGAAC 1120
Db        894 YYYKTRRTRYKTCWWKARWGSWAY 917
```

```
RESULT 10
BX897685/c
LOCUS       BX897685              179286 bp    DNA     linear   VRT 27-MAR-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-234H8, complete sequence.
ACCESSION   BX897685
VERSION     BX897685.6  GI:45772188
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 179286)
  AUTHORS   Barker,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 27, 2004 this sequence version replaced gi:45581072.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-234H8
            is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .179286
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-234H8"
                     /clone_lib="CHORI-211"
ORIGIN
```

```
  Query Match             1.3%;  Score 50.8;  DB 5;  Length 179286;
  Best Local Similarity   44.5%;  Pred. No. 0.18;
  Matches  202;  Conservative    0;  Mismatches  252;  Indels    0;  Gaps    0;

Qy       1003 AAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACT 1062
Db      44434 AAACATTTGAAGAAAAAACGTAATAATAATAATAAAGATAATAATAATAATGATGATGAT 44375

Qy       1063 CTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGT 1122
Db      44374 GATGATGATGATAATAATAATAATAATAGTAATAATAATAATAATAATAGAAATAATAAT 44315

Qy       1123 GATCAGGTTGCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAA 1182
Db      44314 AATGATGATGATGATGATGATAATAATAATAATAATAGTAATAATAATAATAATAATAGA 44255

Qy       1183 TGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCT 1242
Db      44254 AATAATAATAATGATGATGATGATGATGATAATAATAATAATAATAGTAATAATAATAAT 44195

Qy       1243 AAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCC 1302
Db      44194 AATAATAATAATAATAATAAAGATAATAATAATAATAATAATAGAAATAATAACAATGAT 44135

Qy       1303 CAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTA 1362
```

```
Db      44134 GATGATGATGATAATGATGATAATAATAATAGTAATAATAATAATAATAATAATAATAAT 44075

Qy       1363 TCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAA 1422

Db      44074 AATAGTAATAATAATAATAATAATAATAGAAATAATAATAATGATGATGATGATGATGAT 44015

Qy       1423 AATGCCGCTGATCCCTGTGAGGATGATAGAAGTA 1456

Db      44014 AATAATAATAATAATAGTAATAATAATAATATA 43981


RESULT 11
BX248504
LOCUS       BX248504               178247 bp    DNA     linear   VRT 10-JAN-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-72K22 in linkage group 9,
            complete sequence.
ACCESSION   BX248504
VERSION     BX248504.5  GI:40789215
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 178247)
  AUTHORS   Hunter,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 10, 2004 this sequence version replaced gi:38201246.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-72K22
            is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .178247
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-72K22"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match             1.3%;  Score 49.8;  DB 5;  Length 178247;
  Best Local Similarity   44.8%;  Pred. No. 0.33;
  Matches  192;  Conservative    0;  Mismatches  237;  Indels    0;  Gaps    0;

Qy       1344 AAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAG 1403

Db      83879 AATAATGATGATAATAATAATAATAATGATAGTTTATAATAATGATATTAATAATAATAA 83938

Qy       1404 TGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCC 1463

Db      83939 TGATAATAATAATTATATCAATAATAATAATAATGATGATGATAATAATAATAATTATTAA 83998

Qy       1464 GGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGG 1523

Db      83999 TGATAATAATGATATTAATAATGATAATATTAATAGTAATAATAATTGTAATAATAATGA 84058

Qy       1524 GTTAAAATCAAGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATC 1583

Db      84059 TATTATTAATAATAATGATAATAATAATAATGATGATGATGATAATAATAATAATAATAA 84118

Qy       1584 ATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGT 1643

Db      84119 TAGTAATAATAATGATATTCATAATAATAATAATGATGATAATACTGATGATGATGATGA 84178

Qy       1644 TGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCA 1703

Db      84179 TGATGATGATGATAATAATAATAATAGTAATAATAAGGATATTAATAATAATAATAA 84238

Qy       1704 TGATGATGAGAATGATACTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTG 1763

Db      84239 TGATGATAATACTGATGATGATAATAATAATAATAATAATAATAATAATAATAATGATGG 84298

Qy       1764 TCAGCATGT 1772

Db      84299 TGAAAATAT 84307


RESULT 12
CT009752_4
WPCOMMENT
  Sequence split into 9 fragments  LOCUS CT009752 Accession CT009752
     Fragment Name      Begin     End
     CT009752_0             1     110000
     CT009752_1        100001     210000
     CT009752_2        200001     310000
     CT009752_3        300001     410000
     CT009752_4        400001     510000
     CT009752_5        500001     610000
     CT009752_6        600001     710000
     CT009752_7        700001     810000
     CT009752_8        800001     868454
Continuation (5 of 9) of CT009752 from base 400001 (CT009752 Trypanosoma brucei chromos

  Query Match             1.3%;  Score 48.8;  DB 14;  Length 110000;
  Best Local Similarity   41.1%;  Pred. No. 0.6;
  Matches  344;  Conservative    0;  Mismatches  492;  Indels    0;  Gaps    0;

Qy       1367 AAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATG 1426

Db     108986 ATAATAATAATGATAATAATGATAATCATAATAATAATAATAATAATAATAATAATAATA 109045

Qy       1427 CCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGG 1486

Db     109046 ATAATAATAATGATAATAATGATAATAATAATAATCATAATCATGATAATCATAATCATA 109105

Qy       1487 ATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGA 1546

Db     109106 ATAATAATGATAATAATCATAATAATAATAATGATAATGATAATAATAATGATAATAATA 109165
```

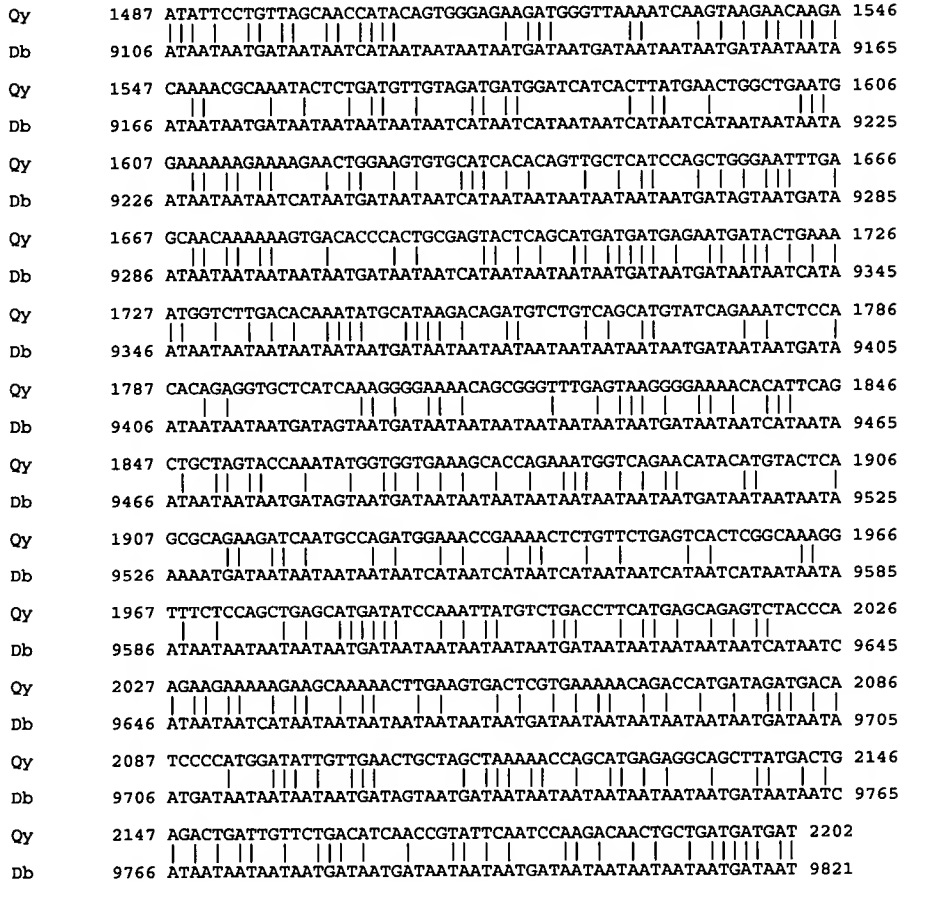


RESULT 13
CT009752_5
WPCOMMENT
Sequence split into 9 fragments LOCUS CT009752 Accession CT009752

Fragment Name	Begin	End
CT009752_0	1	110000
CT009752_1	100001	210000
CT009752_2	200001	310000
CT009752_3	300001	410000
CT009752_4	400001	510000
CT009752_5	500001	610000
CT009752_6	600001	710000
CT009752_7	700001	810000
CT009752_8	800001	868454

Continuation (6 of 9) of CT009752 from base 500001 (CT009752 Trypanosoma brucei chromoso

Query Match 1.3%; Score 48.8; DB 14; Length 110000;
Best Local Similarity 41.1%; Pred. No. 0.6;
Matches 344; Conservative 0; Mismatches 492; Indels 0; Gaps 0;



RESULT 14
BX005453/c
LOCUS BX005453 185541 bp DNA linear VRT 23-AUG-2004
DEFINITION Zebrafish DNA sequence from clone CH211-160B19 in linkage group 19, complete sequence.
ACCESSION BX005453
VERSION BX005453.13 GI:51491109
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 185541)
AUTHORS Bird,C.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Aug 20, 2004 this sequence version replaced gi:51469413.
---------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk

```
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-160E19 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .185541
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-160E19"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match              1.3%;  Score 48.8;  DB 5;  Length 185541;
  Best Local Similarity   45.4%;  Pred. No. 0.6;
  Matches  176;  Conservative    0;  Mismatches  212;  Indels    0;  Gaps    0;

Qy       1372 ATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCT 1431
             |||| | || |  | | |  | |     ||| ||  | ||  | | |||     |
Db     124902 ATAATATCAAATAATAATAATATTAATGATGATGATGATAATAATAATAATAATTATAAT 124843

Qy       1432 GATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATT 1491
              ||     ||| |||||| | ||  |           |  ||    | |   || ||
Db     124842 AATTATAATGATAATGATAATAATAATAATAATAATAAAAATGATAACAAGAATAATGTT 124783

Qy       1492 CCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAA 1551
               |  ||  ||  |||  |  |  || |||     |     | ||| || || | |  |
Db     124782 AATAATAATAATAATAATGATGATGATGATAATAATAGTAATAATAATAATAAAAAATTA 124723

Qy       1552 CGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAA 1611
               ||  |   | |||||    || |||          ||  ||        ||   ||
Db     124722 AAAAATAATAATAATGTTAATAATAATGATAATAATAATAATAATATTAATTATAATAAT 124663

Qy       1612 AAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAAC 1671
              || || | || ||  | ||   || | |  | || | ||      |   |||   |  ||
Db     124662 AATAATAAAAATGATAATGATAATAATAATGATGATGATGATAATAAAAATGATAATAAT 124603

Qy       1672 AAAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGT 1731
              || ||        | |   | || | |  || |||||||  |||||| | | |||  |
Db     124602 AATAATAATGATAATAATAATAATAATAATAATAATAATGATGATGATGATAATAATAAT 124543

Qy       1732 CTTGACACAAATATGCATAAGACAGATG 1759
                | | |  ||||   |||| | | |||
Db     124542 AATAATAATAATAATAATAATAAAAATG 124515
```

```
RESULT 15
AC096088
LOCUS       AC096088              228730 bp    DNA     linear   HTG 13-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-14F23, *** SEQUENCING IN PROGRESS
            ***, 2 unordered pieces.
ACCESSION   AC096088
VERSION     AC096088.6  GI:24941603
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 228730)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 228730)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 228730)
```

```
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 13, 2002 this sequence version replaced gi:22772152.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GEKF
                Center clone name: CH230-14F23
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 220033 bases at least Q40
                Consensus quality: 223217 bases at least Q30
                Consensus quality: 224937 bases at least Q20
                Estimated insert size: 228618; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   227067: contig of 227067 bp in length
            *  227068   227167: gap of unknown length
            *  227168   228730: contig of 1563 bp in length.
FEATURES             Location/Qualifiers
     source          1. .228730
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-14F23"
     misc_feature    219486. .221140
                     /note="wgs_contig"
     misc_feature    222291. .224502
                     /note="wgs_contig"
     misc_feature    224553. .227067
                     /note="wgs_contig"
     gap             227068. .227167
                     /estimated_length=unknown
ORIGIN

  Query Match             1.2%;  Score 48.6;  DB 14;  Length 228730;
  Best Local Similarity   41.6%;  Pred. No. 0.68;
  Matches  117;  Conservative    0;  Mismatches  164;  Indels    0;  Gaps    0;

Qy       1012 GAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCT 1071
              ||||  |      ||  | | ||||| ||| | | | ||   |||||        | |
Db      87894 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 87953

Qy       1072 GAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTT 1131
              |                                                  ||  |||

Db      87954 GNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGAGGAGGAG 88013

Qy       1132 GCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAG 1191
              |  ||| ||   ||   ||  ||  |   |||       |      | ||||      || ||
Db      88014 GAGGAGGAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 88073

Qy       1192 ATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAG 1251
              |      ||  ||  | |   |||||  |      |||  |       ||    |||| ||||
Db      88074 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAAAAGAAGAGAAGAAGAAGAAGAAG 88133

Qy       1252 ACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAA 1292
              | | | ||||||  || ||||||| || ||  | |  |||||
Db      88134 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 88174

Search completed: March  7, 2006, 18:33:28
Job time : 18755 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 12:18:57 ; Search time 2150 Seconds
(without alignments)
12077.043 Million cell updates/sec

Title: US-09-828-068-1
Perfect score: 3896
Sequence: 1 cgcggctgtcggagcaaacg.....gtcaacaccggagaatttac 3896

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3896	100.0	3896	8	ABQ77908	Abq77908 Rice OsEM
	2	3882.4	99.7	4310	6	ABA93084	Aba93084 Oryza sat
	3	2952.2	75.8	9455	6	ABA93085	Aba93085 Oryza sat
	4	173.6	4.5	871	13	ADX09467	Adx09467 Plant ful
	5	73.4	1.9	279	6	ABL73385	Abl73385 Corn tass
c	6	62	1.6	2000	8	ADA71938	Ada71938 Rice gene
	7	53.2	1.4	2000	8	ADA71938	Ada71938 Rice gene
	8	46	1.2	2000	11	ACL37108	Acl37108 Rice stre
	9	45.4	1.2	274	10	ABX85499	Abx85499 Corn ear-
c	10	43	1.1	1788	4	ABL06815	Abl06815 Drosophil
c	11	43	1.1	4773	4	ABL06814	Abl06814 Drosophil
	12	42.8	1.1	574	3	AAC44195	Aac44195 Arabidops
	13	42.2	1.1	110000	6	ABQ69245_18	Continuation (19 o
c	14	42.2	1.1	319630	6	ABQ67194	Abq67194 Listeria
	15	41.4	1.1	246	5	AAS90679	Aas90679 DNA encod
	16	41.4	1.1	1269	6	ABN90734	Abn90734 Staphyloc
	17	41.4	1.1	1269	13	ADS02793	Ads02793 Staphyloc
	18	41.4	1.1	2456	5	AAS75925	Aas75925 DNA encod
c	19	41.4	1.1	2848	4	AAH54196	Aah54196 S. epider
	20	41	1.1	1260	8	ACA46931	Aca46931 Prokaryot
	21	40.8	1.0	2167	3	AAZ32967	Aaz32967 Human wil
c	22	40.6	1.0	297	6	ABN90752	Abn90752 Staphyloc
c	23	40.6	1.0	297	13	ADS02819	Ads02819 Staphyloc
	24	40.6	1.0	2000	11	ACL35887	Acl35887 Rice stre
	25	40.6	1.0	5299	8	AAD49594	Aad49594 Human cyt
	26	40.6	1.0	5322	8	AAD49592	Aad49592 Human cyt
	27	40.6	1.0	5451	10	ADC26234	Adc26234 Human NOV
	28	40.6	1.0	5761	6	ABZ11141	Abz11141 Human pol
	29	40.6	1.0	5761	12	ADM43659	Adm43659 Novel hum
	30	40.4	1.0	686	3	AAA09318	Aaa09318 Human can
c	31	40.4	1.0	96592	10	ADE95860	Ade95860 Human NFK
	32	40.4	1.0	110000	6	ABA03041_17	Continuation (18 o
c	33	40.2	1.0	2000	11	ACL37108	Acl37108 Rice stre
	34	40.2	1.0	15518	6	ABL34172	Abl34172 Human imm
	35	40.2	1.0	15518	6	ABL34624	Abl34624 Human met
	36	40.2	1.0	15518	6	ABL70607	Abl70607 Chemicall
	37	40.2	1.0	15518	7	ADS99885	Ads99885 Bisulphit
	38	40	1.0	412	9	ACH35151	Ach35151 Human end
	39	40	1.0	3136	14	ADZ14609	Adz14609 DNA C14or
c	40	39.8	1.0	15122	6	ABL32775	Abl32775 Human imm
c	41	39.6	1.0	882	13	ADS64035	Ads64035 Bacterial
	42	39.6	1.0	3257	4	AAH54703	Aah54703 S. epider
c	43	39.6	1.0	3836	14	ADW18048	Adw18048 Pinus rad
c	44	39.4	1.0	843	5	AAS92080	Aas92080 DNA encod
	45	39.2	1.0	1251	13	ADO82093	Ado82093 Plant ful

ALIGNMENTS

RESULT 1
ABQ77908
ID ABQ77908 standard; cDNA; 3896 BP.
XX
AC ABQ77908;
XX
DT 24-JAN-2003 (first entry)
XX
DE Rice OsEMF1-encoding cDNA, SEQ ID NO:1.
XX
KW Rice; OsEMF1; EMF; embryonic flower; plant; floral repressor;
KW reproductive development; flower development; transgenic plant;
KW antisense suppression; transgenic; reciprocal negative interaction;
KW flower meristem identity gene; flowering time; shoot development;
KW seed yield; agriculture; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 241. .3414
FT /*tag= a
FT /product= "OsEMF1 protein"
XX
PN WO200280659-A1.
XX
PD 17-OCT-2002.
XX
PF 05-APR-2002; 2002WO-US012675.
XX
PR 06-APR-2001; 2001US-00828068.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Moon Y, Chen L, Sung ZR;
XX
DR WPI; 2003-046831/04.
DR P-PSDB; ABB99878.
XX
PT New isolated OsEMF1 nucleic acid molecule and encoded polypeptide, useful
PT for plant genetic engineering, in particular controlling reproductive
PT development in rice.
XX

PS Claim 3; Fig 1; 47pp; English.
XX
CC The invention relates to a rice EMF (embryonic flower) gene designated
CC OsEMF1 (cDNA given in ABQ77908) and its encoded protein (ABB99878). The
CC OsEMF1 protein has a molecular weight of 116.4 kD and exhibits 37%
CC homology and 20% identity with Arabidopsis thaliana EMF1 protein
CC (ABB99879). OsEMF1, like other EMF gene products, acts as a floral
CC repressor, suppressing the transition from vegetative growth to
CC reproductive development. It also delays the inflorescence to flower
CC transition, indicating that there is a reciprocal negative interaction
CC between OsEMF and flower meristem identity genes. The invention also
CC encompasses a transgenic plant comprising an expression cassette
CC containing an OsEMF1 nucleic acid (particularly in an antisense
CC orientation) under the control of a plant promoter. OsEMF1 nucleic acids
CC and proteins may be used for controlling reproductive development in
CC plants, particularly monocotyledonous plants and especially rice. In
CC particular, OsEMF1 nucleic acids and proteins may be used to control
CC flowering time, shoot development and seed yield. Controlling or
CC inhibiting the expression of genes which mediate these processes enables
CC new varieties of rice with different flowering times and seed yield to be
CC developed. The present sequence represents OsEMF1 cDNA. Note: The present
CC sequence is described as SEQ ID NO:1 in the claims and sequence listing,
CC but the sequence referred to as SEQ ID NO:1 in the examples (not shown in
CC the specification) is described as an Arabidopsis thaliana EMF1 genomic
CC clone
XX
SQ Sequence 3896 BP; 1176 A; 791 C; 935 G; 994 T; 0 U; 0 Other;

Query Match 100.0%; Score 3896; DB 8; Length 3896;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

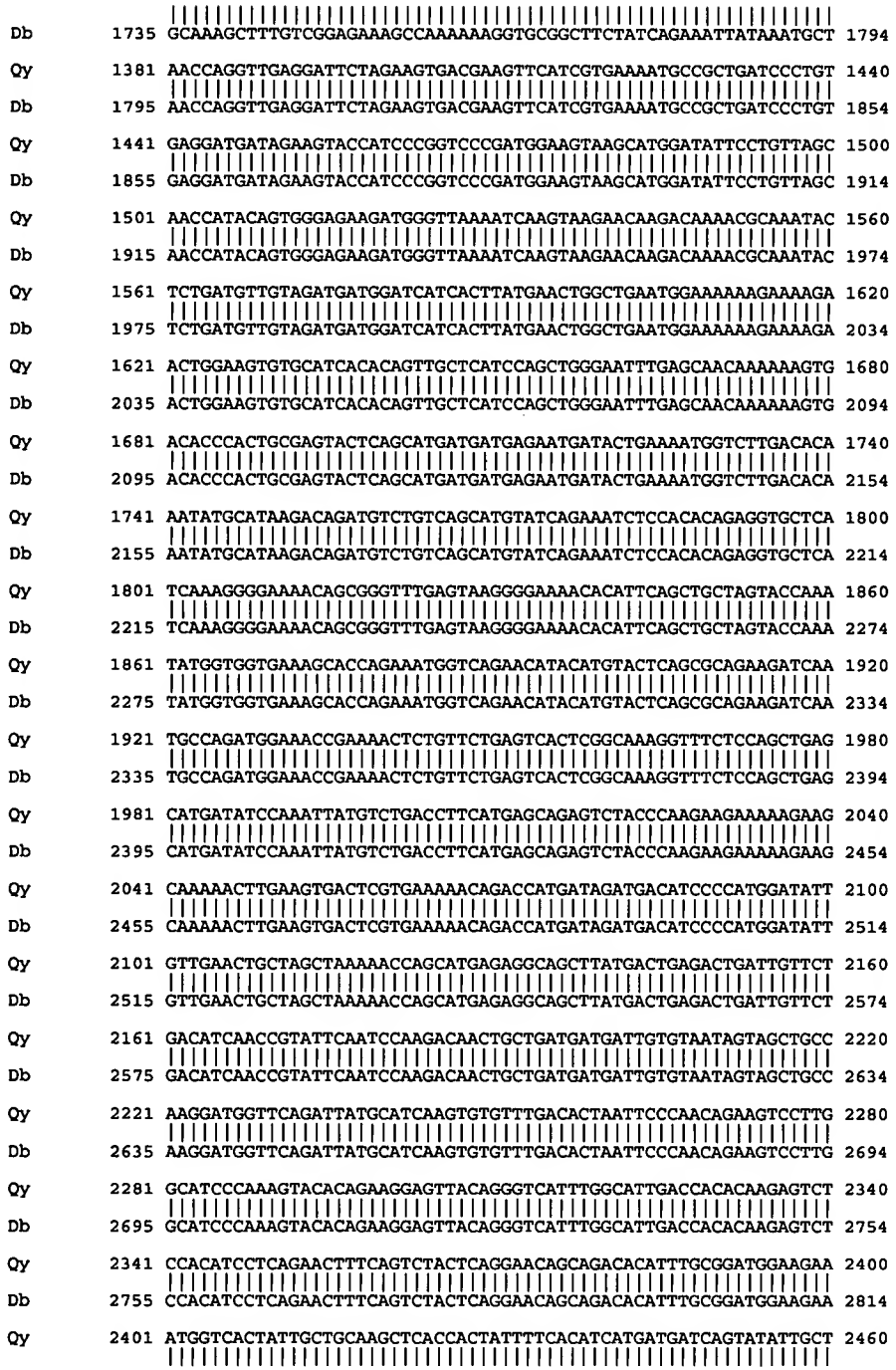
```
Qy          1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60

Qy         61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120

Qy        121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180

Qy        181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240

Qy        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300

Qy        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360

Qy        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420

Qy        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480

Qy        481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540

Qy        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600

Qy        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660

Qy        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720

Qy        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780

Qy        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840

Qy        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900

Qy        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960

Qy        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020

Qy       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080

Qy       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140

Qy       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200

Qy       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260

Qy       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320

Qy       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380

Qy       1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440

Qy       1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500

Qy       1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560

Qy       1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620

Qy       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680

Qy       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
Qy       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
Db       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
Qy       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
Db       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
Qy       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
Db       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
Qy       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
Db       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
Qy       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
Db       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
Qy       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
Db       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
Qy       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
Db       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
Qy       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
Db       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
Qy       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
Db       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
Qy       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
Db       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
Qy       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
Db       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
Qy       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
Db       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
Qy       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
Db       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
Qy       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
Db       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
Qy       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
Db       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
Qy       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
Db       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
Qy       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
Db       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
Qy       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
Db       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
Qy       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
Db       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
Qy       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
Db       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
Qy       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
Db       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
Qy       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
Db       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
Qy       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
Db       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
Qy       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
Db       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
Qy       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
Db       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
Qy       3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
Db       3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
Qy       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
Db       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
Qy       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3420
Db       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3420
Qy       3421 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3480
Db       3421 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3480
Qy       3481 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3540
Db       3481 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3540
Qy       3541 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 3600
Db       3541 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 3600
Qy       3601 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 3660
Db       3601 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 3660
Qy       3661 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTC 3720
Db       3661 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTC 3720
Qy       3721 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 3780
Db       3721 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 3780
Qy       3781 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 3840
Db       3781 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 3840
Qy       3841 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
Db       3841 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
```

RESULT 2
ABA93084
ID ABA93084 standard; cDNA; 4310 BP.
XX
AC ABA93084;
XX
DT 11-APR-2002 (first entry)
XX
DE Oryza sativa brassinosteroid response related protein encoding cDNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
KW signal transfer system; brassinosteroid hormone; growth promotion;
KW increased yield; quality improvement; ripeness promotion;
KW stress relaxation; chemical resistance; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 655. .3828
FT /*tag= a
FT /product= "brassinosteroid response related protein"
XX
PN JP2001327287-A.
XX
PD 27-NOV-2001.
XX
PF 19-MAY-2000; 2000JP-00149106.
XX
PR 19-MAY-2000; 2000JP-00149106.
XX
PA (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
PA (SEIB-) SEIBUTSUKEI TOKUTEI SANGYO GIJUTSU.
XX
DR WPI; 2002-135739/18.
DR P-PSDB; ABB05428.
XX
PT New gene involved in brassinosteroid responses useful for controlling the
PT effects such as growth promotion, increased yield, quality improvement,
PT ripeness promotion, stress relaxation and chemical resistance.
XX
PS Disclosure; Page 6-10; 19pp; Japanese.
XX
CC The present invention describes a polynucleotide encoding a plant gene
CC which can control the signal transfer system of brassinosteroid hormone.
CC The polynucleotide can be used for controlling the effects such as growth
CC promotion, increased yield, quality improvement, ripeness promotion,
CC stress relaxation and chemical resistance. The present sequence encodes a
CC brassinosteroid response related protein isolated from rice (Oryza
CC sativa), from the present invention
XX
SQ Sequence 4310 BP; 1237 A; 986 C; 1022 G; 1065 T; 0 U; 0 Other;

Query Match 99.7%; Score 3882.4; DB 6; Length 4310;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3894; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy          1 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CGCGGCTGTCGGAGCAAACGCAAACCCCCCAGGTTGTTCTAGCGTGTGCAGCGGCTAGCT 474

Qy         61 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTGGTT 534

Qy        121 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATAT 594

Qy        181 TGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 240
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        595 TGAGTTGTGCCTCGGCTGTGCTAGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGAT 654

Qy        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 714

Qy        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 774

Qy        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 834

Qy        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 894

Qy        481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 954

Qy        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1014

Qy        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1074

Qy        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1134

Qy        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1194

Qy        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1254

Qy        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1314

Qy        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1374

Qy        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1434

Qy       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1494

Qy       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1554

Qy       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1614

Qy       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1674

Qy       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1734

Qy       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380
```


Db 1735 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1794
Qy 1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440
Db 1795 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1854
Qy 1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500
Db 1855 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1914
Qy 1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560
Db 1915 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1974
Qy 1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620
Db 1975 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2034
Qy 1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
Db 2035 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2094
Qy 1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
Db 2095 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2154
Qy 1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
Db 2155 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2214
Qy 1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
Db 2215 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2274
Qy 1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
Db 2275 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2334
Qy 1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
Db 2335 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2394
Qy 1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
Db 2395 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2454
Qy 2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
Db 2455 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2514
Qy 2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
Db 2515 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2574
Qy 2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
Db 2575 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2634
Qy 2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
Db 2635 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2694
Qy 2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
Db 2695 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2754
Qy 2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
Db 2755 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2814
Qy 2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
Db 2815 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2874
Qy 2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
Db 2875 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2934
Qy 2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
Db 2935 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2994
Qy 2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
Db 2995 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3054
Qy 2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
Db 3055 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3114
Qy 2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
Db 3115 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3174
Qy 2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
Db 3175 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3234
Qy 2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
Db 3235 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3294
Qy 2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
Db 3295 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3354
Qy 2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
Db 3355 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3414
Qy 3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
Db 3415 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3474
Qy 3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
Db 3475 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3534
Qy 3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
Db 3535 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3594
Qy 3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
Db 3595 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3654
Qy 3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
Db 3655 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3714
Qy 3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
Db 3715 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3774
Qy 3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3420
Db 3775 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGA 3834
Qy 3421 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3480
Db 3835 AAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTTATGGCGCCAAA 3894
Qy 3481 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3540
Db 3895 TAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTG 3954

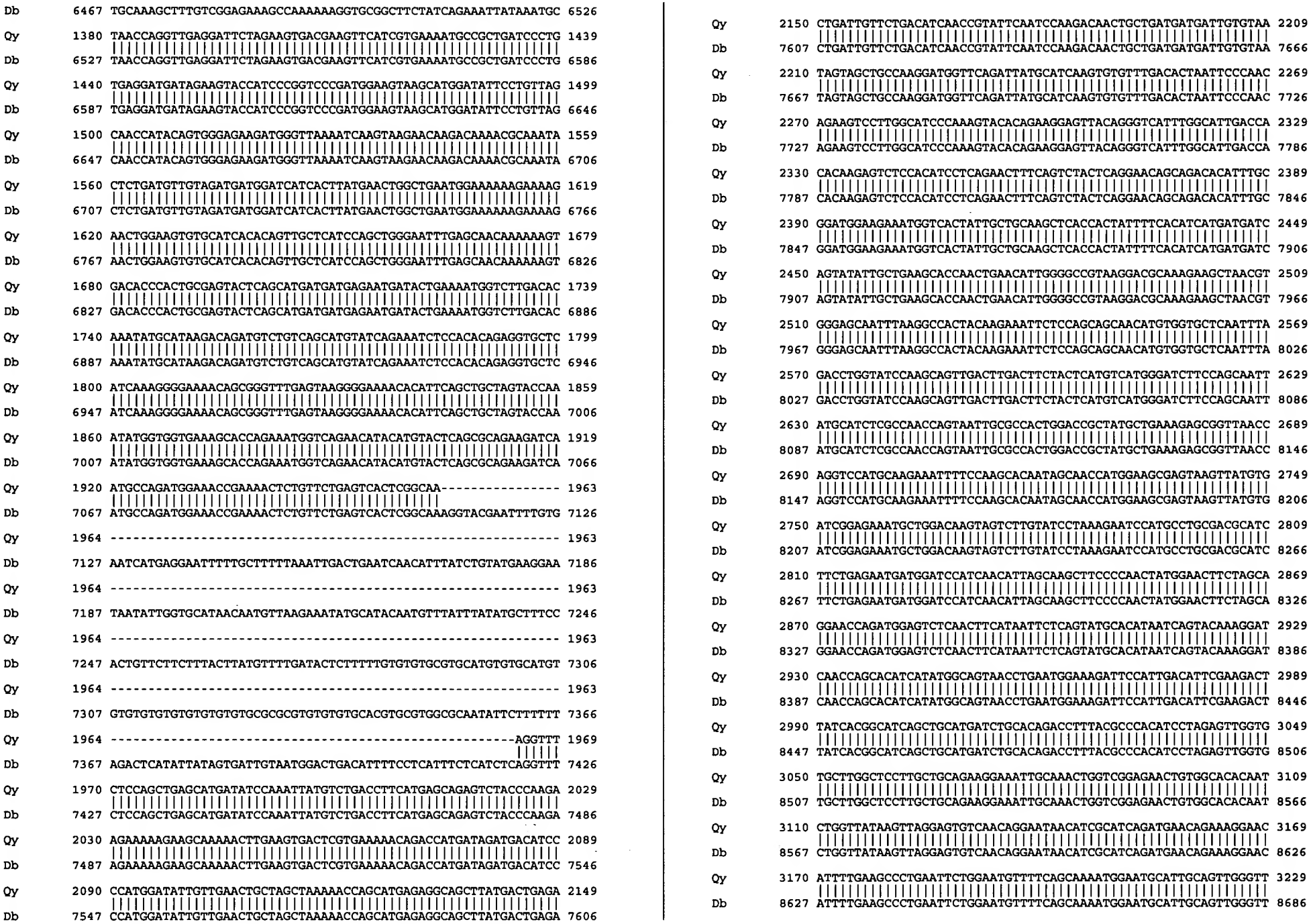
```
Qy       3541 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3955 GATGCCCCATTTTCTGGATGGGAACCTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCA 4014

Qy       3601 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4015 GCATCCGGTTTTTGTTTTGCCAGTCCAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACT 4074

Qy       3661 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTC 3720
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       4075 CATACTAGTGCGCTTGTTTGTACAAGGAGAAATGTGTAACCTTGTTG-AAAAAATGTCTC 4133

Qy       3721 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4134 CCCCATTTTGTAATTACCATAAGGAGGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGG 4193

Qy       3781 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4194 CGAGAAATGGTTTTGTCGGTGTTAAGGTTGAAACGACTAGCTCTCGTTATCAATGTGTTG 4253

Qy       3841 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4254 TAAACTTCTAGATTGATGTGTTACCTTACTCTTGAAGTCAACACCGGAGAATTTAC 4309
```

RESULT 3
ABA93085
ID ABA93085 standard; DNA; 9455 BP.
XX
AC ABA93085;
XX
DT 11-APR-2002 (first entry)
XX
DE Oryza sativa brassinosteroid response related protein related DNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
KW signal transfer system; brassinosteroid hormone; growth promotion;
KW increased yield; quality improvement; ripeness promotion;
KW stress relaxation; chemical resistance; gene; ds.
XX
OS Oryza sativa.
XX
PN JP2001327287-A.
XX
PD 27-NOV-2001.
XX
PF 19-MAY-2000; 2000JP-00149106.
XX
PR 19-MAY-2000; 2000JP-00149106.
XX
PA (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
PA (SEIB-) SEIBUTSUKEI TOKUTEI SANGYO GIJUTSU.
XX
DR WPI; 2002-135739/18.
XX
PT New gene involved in brassinosteroid responses useful for controlling the
PT effects such as growth promotion, increased yield, quality improvement,
PT ripeness promotion, stress relaxation and chemical resistance.
XX
PS Disclosure; Page 13-16; 19pp; Japanese.
XX
CC The present invention describes a polynucleotide encoding a plant gene
CC which can control the signal transfer system of brassinosteroid hormone.
CC The polynucleotide can be used for controlling the effects such as growth
CC promotion, increased yield, quality improvement, ripeness promotion,
CC stress relaxation and chemical resistance. The present sequence
CC represents a DNA sequence related to a brassinosteroid response related
CC protein isolated from rice (Oryza sativa), from the present invention
XX
SQ Sequence 9455 BP; 2571 A; 2001 C; 2090 G; 2793 T; 0 U; 0 Other;

Query Match 75.8%; Score 2952.2; DB 6; Length 9455;

Best Local Similarity 87.4%; Pred. No. 0;
Matches 3487; Conservative 0; Mismatches 23; Indels 481; Gaps 4;

```
Qy        386 ACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGAGGGTATGTTGCTCTTCTTC 445
              | | | |  | ||    |||  |  ||  |  | ||||||||||||||||||||||
Db       5447 ATGCATCTTTATAATGATAATCTTATTTTTCTTGTACCAGAGGGTATGTTGCTCTTCTTC 5506

Qy        446 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5507 AGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTG 5566

Qy        506 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5567 ATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCT 5626

Qy        566 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5627 CGAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCG 5686

Qy        626 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5687 CAAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGC 5746

Qy        686 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5747 CTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATG 5806

Qy        746 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5807 CTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGC 5866

Qy        806 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA---------- 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5867 CTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGGTATGGTAG 5926

Qy        855 ---------------------------------------------------------- 854

Db       5927 ATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGAATAAACTAATGT 5986

Qy        855 ---------------AGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 899
                              ||||||||||||||||||||||||||||||||||||||||||||
Db       5987 TTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGT 6046

Qy        900 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6047 CTCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATG 6106

Qy        960 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6107 GCACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGT 6166

Qy       1020 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6167 CCTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTG 6226

Qy       1080 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6227 CAATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCA 6286

Qy       1140 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6287 GTGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAA 6346

Qy       1200 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 1259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6347 TGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAA 6406

Qy       1260 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 1319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6407 GAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGA 6466

Qy       1320 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 1379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 6467 TGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGC 6526
Qy 1380 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 1439
Db 6527 TAACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTG 6586
Qy 1440 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 1499
Db 6587 TGAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAG 6646
Qy 1500 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 1559
Db 6647 CAACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATA 6706
Qy 1560 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 1619
Db 6707 CTCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG 6766
Qy 1620 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 1679
Db 6767 AACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGT 6826
Qy 1680 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 1739
Db 6827 GACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACAC 6886
Qy 1740 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 1799
Db 6887 AAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTC 6946
Qy 1800 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 1859
Db 6947 ATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAA 7006
Qy 1860 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 1919
Db 7007 ATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCA 7066
Qy 1920 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAA---------- 1963
Db 7067 ATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTACGAATTTTGTG 7126
Qy 1964 ---------- 1963
Db 7127 AATCATGAGGAATTTTTGCTTTTTAAATTGACTGAATCAACATTTATCTGTATGAAGGAA 7186
Qy 1964 ---------- 1963
Db 7187 TAATATTGGTGCATAACAATGTTAAGAAATATGCATACAATGTTTATTTATATGCTTTCC 7246
Qy 1964 ---------- 1963
Db 7247 ACTGTTCTTCTTTACTTATGTTTTGATACTCTTTTTGTGTGTGCGTGCATGTGTGCATGT 7306
Qy 1964 ---------- 1963
Db 7307 GTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGCACGTGCGTGGCGCAATATTCTTTTTT 7366
Qy 1964 ----------AGGTTT 1969
Db 7367 AGACTCATATTATAGTGATTGTAATGGACTGACATTTTCCTCATTTCTCATCTCAGGTTT 7426
Qy 1970 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 2029
Db 7427 CTCCAGCTGAGCATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGA 7486
Qy 2030 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 2089
Db 7487 AGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCC 7546
Qy 2090 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 2149
Db 7547 CCATGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGA 7606
Qy 2150 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 2209
Db 7607 CTGATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAA 7666
Qy 2210 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 2269
Db 7667 TAGTAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAAC 7726
Qy 2270 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 2329
Db 7727 AGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCA 7786
Qy 2330 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 2389
Db 7787 CACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGC 7846
Qy 2390 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 2449
Db 7847 GGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATC 7906
Qy 2450 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 2509
Db 7907 AGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGT 7966
Qy 2510 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 2569
Db 7967 GGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTA 8026
Qy 2570 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 2629
Db 8027 GACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATT 8086
Qy 2630 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 2689
Db 8087 ATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACC 8146
Qy 2690 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 2749
Db 8147 AGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTG 8206
Qy 2750 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 2809
Db 8207 ATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATC 8266
Qy 2810 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 2869
Db 8267 TTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCA 8326
Qy 2870 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 2929
Db 8327 GGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGAT 8386
Qy 2930 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 2989
Db 8387 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 8446
Qy 2990 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 3049
Db 8447 TATCACGGCATCAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTG 8506
Qy 3050 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 3109
Db 8507 TGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAAT 8566
Qy 3110 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 3169
Db 8567 CTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAAC 8626
Qy 3170 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 3229
Db 8627 ATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTT 8686

```
Qy       3230 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 3289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8687 CTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCA 8746

Qy       3290 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 3349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8747 GAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAA 8806

Qy       3350 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 3409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8807 CTAACAAGAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACC 8866

Qy       3410 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 3469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8867 GCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACATTTACAATCATACATTTCATCTTT 8926

Qy       3470 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTA-------- 3521
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8927 ATGGCGCCAAATAGTCATACTGTAAGAGGAGGGCTTTGCTGGATCTGCTGTAAGGTAAGT 8986

Qy       3522 ------------------------------------------------------------ 3521

Db       8987 TGAACTTTTTCTTCTTGCAAGTTTATCAGTTTAAGAAAAAAGAATGATTACTTATGTTAG 9046

Qy       3522 ----------------AGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 3565
                               ||||||||||||||||||||||||||||||||||||||||||||
Db       9047 CAAGGATGGTTCTTGCAGGCTTCTTGTAAGTTGTGGATGCCCCATTTTCTGGATGGGAAC 9106

Qy       3566 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 3625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9107 CTGCCAGACAGTGAACAAGGGCTTTGCAAGGTGCAGCATCCGGTTTTTGTTTTGCCAGTC 9166

Qy       3626 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 3685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9167 CAAGAAACGTCCTCCTGTTACTTTGTAGTTGTACTCATACTAGTGCGCTTGTTTGTACAA 9226

Qy       3686 GGAGAAATGTGTAACCTTGTTGAAAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 3745
             ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       9227 GGAGAAATGTGTAACCTTGTTG-AAAAAATGTCTCCCCCATTTTGTAATTACCATAAGGA 9285

Qy       3746 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 3805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9286 GGTTTATAGTGTTGTGAGCTGTGTGTGACTGACGGCGAGAAATGGTTTTGTCGGTGTTAA 9345

Qy       3806 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 3865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9346 GGTTGAAACGACTAGCTCTCGTTATCAATGTGTTGTAAACTTCTAGATTGATGTGTTACC 9405

Qy       3866 TTACTCTTGAAGTCAACACCGGAGAATTTAC 3896
             ||||||||||||||||||||||||||||||
Db       9406 TTACTCTTGAAGTCAACACCGGAGAATTTAC 9436
```

RESULT 4
ADX09467
ID ADX09467 standard; cDNA; 871 BP.
XX
AC ADX09467;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 4042.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 4042; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 871 BP; 225 A; 197 C; 230 G; 219 T; 0 U; 0 Other;

Query Match 4.5%; Score 173.6; DB 13; Length 871;
Best Local Similarity 65.3%; Pred. No. 3.1e-39;
Matches 311; Conservative 0; Mismatches 144; Indels 21; Gaps 3;

```
Qy       2960 ATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCATCA---GCTGCATGATCTGC 3016
              ||||||     ||  ||||| | |||||||| || |||| |||    | ||| ||  |||
Db        105 ATGGAAGCCAACCGCTGACACTGGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGC 164

Qy       3017 ACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAA 3076
              ||| |||||||||||| || ||| | || |||||||| || || |||||||||| |||| |
Db        165 GCAGGCCTTTACGCCCTCACCCTCGTGTCGGTGTCTCGGTTCATTGCTGCAGCAGGAGA 224

Qy       3077 TTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAG 3136
              | |||||||||| | |||||| ||  | || ||||| || |   |||| |  |  | ||
Db        225 TCGCAAACTGGTCTGGGAACTGCGGGCCGCAGTCTGGGTACAGACTAGGTGATTGTAAAG 284

Qy       3137 GAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGT 3196
              | | ||   ||| | |    ||||||   | | | |  ||  || |||
Db        285 GGACAAGGCCGCCTGATGCCAACAGAGCAGGAAACTACGAGACCTTGA------------ 332

Qy       3197 TTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCT------CCAGTGCAGATT 3250
              || |||  ||||||  |  ||||||||||||||||||||      ||| |   || |
Db        333 GCTCGGCAGGATGGAACACCCTGCAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGT 392
```

```
Qy       3251 TTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTC 3310
              |   |  || |  | |  | |||  ||| |||||| ||||| || |  |    |||
Db        393 ACCGGTTTCCATGGTATGGTACGGGTCAGCCTTCGACCAGTGGCAATGGGAGGACCATTC 452

Qy       3311 ATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATT 3370
              ||||  ||||||  | | |||||  |||||||||||| | |||||||  |||||||| ||||
Db        453 ATCCGTTGGATAAGCTCGTGAGGAAGGATATCTGTGTGACTAACAGAAACCCAGCCGATT 512

Qy       3371 TTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGAAGCAGAAAGTGG 3426
              | |||  ||| ||| || | |||||||| |||     ||  ||||  ||| |  ||
Db        513 TCACTGTAATTAGTGACAAGAACGAGTACATGAGAACCCTTTGAAAGAGAGAAAGG 568
```

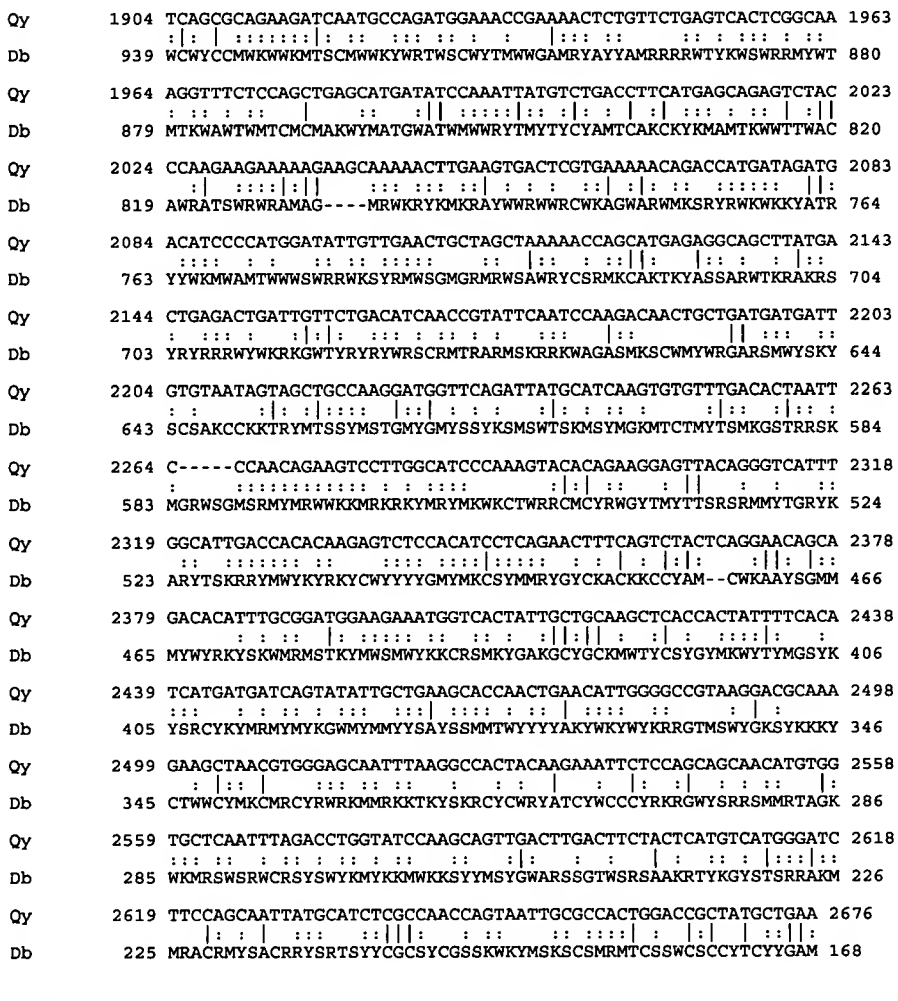
RESULT 5
ABL73385
ID ABL73385 standard; cDNA; 279 BP.
XX
AC ABL73385;
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:2759.
XX
KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS Zea mays.
XX
PN US2001051335-A1.
XX
PD 13-DEC-2001.
XX
PF 16-APR-1999; 99US-00294093.
XX
PR 21-APR-1998; 98US-0082567P.
XX
PA (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
DR WPI; 2002-163647/21.
XX
PT Novel purified corn tassel-derived polynucleotide useful for determining
PT altered gene expression, to recover regulatory elements and to follow
PT inheritance of desirable characteristics through hybrid breeding
PT programs.
XX
PS Claim 1; SEQ ID NO 2759; 201pp; English.
XX
CC The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful in
CC the evaluation, and alteration of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality and yield, and as molecular markers for studying inheritance of
CC multigene traits in a plant breeding program. (I) can be used to produce
CC a tassel-specific profile of gene transcription, a transcript image, to
CC clone regulatory elements for use in transformation vectors, to express a
CC polypeptide, to identify, isolate or extend identical or related corn
CC tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences
XX
SQ Sequence 279 BP; 70 A; 74 C; 73 G; 56 T; 0 U; 6 Other;

Query Match 1.9%; Score 73.4; DB 6; Length 279;
Best Local Similarity 71.3%; Pred. No. 2.8e-10;
Matches 122; Conservative 0; Mismatches 45; Indels 4; Gaps 2;

```
Qy       2960 ATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCATCA---GCTGCATGATCTGC 3016
              ||||||     ||  ||| | | |||||||| || |||| |||     | ||| ||  |||
Db        105 ATGGAAGCCAACCGCTGANACTGGAAGACTTGTCTCGGCGTCATTNCCAGCAAGACTTGC 164

Qy       3017 ACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAA 3076
              ||| |||||||||| | || ||| || | |||||||| || || |||||||||| |||| |
Db        165 GCAGGCCTTTACGCNCTCACCCTCGANTCGGTGTGCTCGGTTCATTGCTGCAGCAGGAGA 224

Qy       3077 TTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAG 3127
              | |||||||||||  | ||| || ||  | || ||||| || |   |||| |
Db        225 TCGCAAACTGGTCTGGGAA-TGCGGGCCGCAGTCTGGGTANAGACTAGGTG 274
```

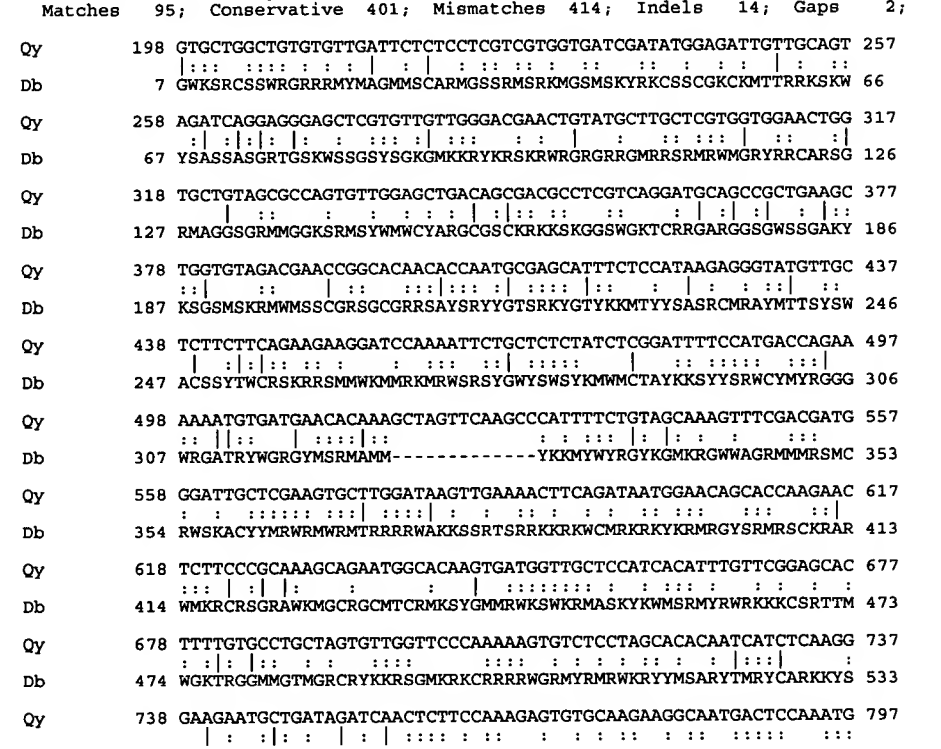
RESULT 6
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 1.6%; Score 62; DB 8; Length 2000;
Best Local Similarity 10.5%; Pred. No. 2e-06;
Matches 82; Conservative 360; Mismatches 325; Indels 11; Gaps 3;



RESULT 7
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 1.4%; Score 53.2; DB 8; Length 2000;
Best Local Similarity 10.3%; Pred. No. 0.00075;
Matches 95; Conservative 401; Mismatches 414; Indels 14; Gaps 2;



```
Db        534 YSAARKARCWYRGKGYYWAGMWMKRYKRMYMYKMMWWYKRKYSKCSWYCKMSYYASCMKS 593

Qy        798 CAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGA 857

Db        594 ARKAGAKMCKRSKMSAWSKSMRSSRKCRKCASKRSSAKRYAMMGGMTSGSRMSRWKSYTC 653

Qy        858 TTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTC 917

Db        654 YWRKWGSMKSTCTWMYYMSKYTYAKYGSYWRYRYRAWCMYMWRWYYYRYRSYMTYMAWYT 713

Qy        918 TGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGG 977

Db        714 SSTRMAMTGMKYSGRYWTSWYKYCKCSWKYRSMWYYWSWWWAKTWKMWRRYATRMMWMWY 773

Qy        978 TGCAGATCAACCTCCATCCACTCCAAAACTTTCTG-AAGTGGTCCTCAAAAGAAATGAAG 1036

Db        774 RYSMKWYTWCTMWGYWWYWWRTYMKMRYMWYKCTKTYWYWSATYWTGTWAAWWMAKTKMR 833

Qy       1037 ATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATC 1096

Db        834 MGMTGAKTRGRARKARYWWKWATWCATKRWMTKGKGAKWAWTWMAKAWRKYYWSWMRAWY 893

Qy       1097 CTAACCCAATGTCTGGAAAGGAAC 1120

Db        894 YYYKTRRTRYKTCWWKARWGSWAY 917


RESULT 8
ACL37108
ID   ACL37108 standard; cDNA; 2000 BP.
XX
AC   ACL37108;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Rice stress-regulated promoter SEQ ID NO:15671.
XX
KW   ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW   agriculture.
XX
OS   Oryza sativa.
XX
PN   WO2003008540-A2.
XX
PD   30-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019668.
XX
PR   22-JUN-2001; 2001US-0300112P.
PR   24-AUG-2001; 2001US-0314662P.
PR   26-SEP-2001; 2001US-0325277P.
PR   21-NOV-2001; 2001US-0332132P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
PT   New stress-responsive nucleic acid, useful for altering the
PT   responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT   stress, salt stress or osmotic stress.
XX
PS   Claim 48; SEQ ID NO 15671; 89pp; English.
XX
CC   The invention relates to novel abiotic stress responsive polynucleotides
CC   and polypeptides. Also disclosed are vectors, expression cassettes, host
CC   cells, and plants containing such polynucleotides. Also disclosed are
CC   methods for using the polynucleotides and polypeptides to alter the
CC   responsiveness of a plant to abiotic stress. The invention is useful in
CC   agriculture. The nucleic acid is useful for determining whether a test
CC   plant has been exposed to an abiotic stress condition. It is also useful
CC   for selecting an agent that alters abiotic stress regulated
CC   polynucleotide expression in a plant cell, and to identify a homolog or
CC   ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC   molecule and the polypeptide encoded by it are useful in altering the
CC   responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC   stress, osmotic stress or any of their combinations. The present sequence
CC   is used in the exemplification of the invention
XX
SQ   Sequence 2000 BP; 510 A; 350 C; 268 G; 518 T; 0 U; 354 Other;

  Query Match             1.2%;  Score 46;  DB 11;  Length 2000;
  Best Local Similarity   10.4%;  Pred. No. 0.095;
  Matches   46;  Conservative  217;  Mismatches  172;  Indels    7;  Gaps    2;

Qy       1548 AAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTA--TGAACTGGCTGAAT 1605

Db         21 MMARRRSAMYKWSYYYSRTKYYKWYYYRTCSYKRWWMTSMMKYRAYMMRWYWKKMWAMWY 80

Qy       1606 GGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTG 1665

Db         81 AWCMWYAYGYMMWKWWYSRTMTYTRAYYKKWKMWRRKMSRMAAAWYYWSYWRTSRWKWTG 140

Qy       1666 AGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCATGA-----TGATGAGAATGATA 1720

Db        141 YRKKRWWRARRYMYWTYWCYTWSWMRWYRYWKMMKYRAYKWMAYTTTYMTKCCMWWWCTY 200

Qy       1721 CTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAA 1780

Db        201 GYKWKWAAWRASAKRARRAWMMMRWKSMWTWAATWSWWSMMATWKRKMRWGWMWAYSSTW 260

Qy       1781 TCTCCACACAGAGGTGCTCATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACAC 1840

Db        261 MMRWWWYCCTRRAMYKMKAATMTRCTWYMWMAAMTSAWRAWCAGWATSYATKWYAWRTTY 320

Qy       1841 ATTCAGCTGCTAGTACCAAATATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATG 1900

Db        321 MKYRRRYYACMAWTRWTKMWSMMRTKKWWYTMARSWGCKKWMWAKARMATMCMAYTYSTY 380

Qy       1901 TACTCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGG 1960

Db        381 RYYAWRRRWMRWSWAAWYSAKYWYAKSYYYMYKCARKWWTTYAYYAAAAKATRTAAYMGR 440

Qy       1961 CAAAGGTTTCTCCAGCTGAGCA 1982

Db        441 TWYTMCWTMYTKYMKCMRWWYM 462


RESULT 9
ABX85499
ID   ABX85499 standard; cDNA; 274 BP.
XX
AC   ABX85499;
XX
DT   24-APR-2003  (first entry)
XX
DE   Corn ear-derived polynucleotide (cpd) #3959.
XX
KW   Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW   structural gene; functional gene; regulatory gene;
KW   corn ear-specific profile; gene transcription; gene expression;
KW   hybrid plant; desirable trait expression; plant breeding program;
KW   inheritance; desired characteristic; growth; development;
KW   disease resistance; environmental adaptability; quality; yield;
KW   multigene trait; plant; gene; ss.
XX
OS   Zea mays.
XX
PN   US6476212-B1.
XX
PD   05-NOV-2002.
XX
PF   14-MAY-1999;   99US-00313294.
XX
```

```
PR   26-MAY-1998;   98US-0086722P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lalgudi RV,  Ito LY,  Sherman BK;
XX
DR   WPI; 2003-208840/20.
XX
PT   Novel purified corn-ear derived polynucleotide useful as hybridization
PT   probe for detecting polynucleotide in sample, and for identifying,
PT   evaluating, and altering desired characteristics associated with growth,
PT   development.
XX
PS   Example; SEQ ID NO 3959; 390pp; English.
XX
CC   The present invention relates to the isolation of corn ear-derived
CC   polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC   and SATMON023. Some of the cpds uniquely identify structural, functional,
CC   and regulatory genes of corn ear. The polynucleotides sequences are
CC   useful for detecting cpds in a sample, for producing a corn ear-specific
CC   profile of gene transcription, for detecting altered gene expression in
CC   inbred or hybrid plants, and for screening several molecules for specific
CC   binding to the polynucleotide. The cdps are useful to identify, isolate,
CC   or extend identical or related corn-ear nucleic acid sequences from DNA
CC   libraries, and in nucleic acid amplification or hybridisation techniques
CC   to follow the expression of desirable traits through plant breeding
CC   programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC   follow the inheritance of desired characteristics associated with growth
CC   and development, disease resistance, environmental adaptability, quality,
CC   and yield of corn. The cdps are also useful as molecular markers for
CC   studying inheritance and multigene traits in a plant breeding program.
CC   The cpds are useful for producing purified corn-ear polypeptides by
CC   recombinant techniques. They are also useful in diagnostic assays to
CC   detect or confirm conditions or diseases associated with abnormal levels
CC   of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC   polynucleotides (cpds) of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ   Sequence 274 BP; 67 A; 75 C; 75 G; 56 T; 0 U; 1 Other;

  Query Match             1.2%;  Score 45.4;  DB 10;  Length 274;
  Best Local Similarity   68.2%;  Pred. No. 0.042;
  Matches  107;  Conservative    0;  Mismatches   46;  Indels    4;  Gaps    3;

Qy       2960 ATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGA--TCTGCA 3017
              ||||||     ||  |||||  | ||||||||  || ||||  |||   | ||  |    ||
Db        105 ATGGAAGCCAACCGCTGACACTGGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGC 164

Qy       3018 CAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAAT 3077
              ||||||||||| || ||| | || |||||||| || ||  |||||||| |||| ||
Db        165 GCAGCCTTTACGCCCTCACCCTCGTGTCGGTGTGCTCGGTTCA-TGCTGCAGCAGGAGAT 223

Qy       3078 TGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGT 3114
              |||||  |||||  | ||||||  ||   | ||     |||
Db        224 CGCAAA-TGGTCTGGGAACTGCGGGCCGCATCTGGGT 259


RESULT 10
ABL06815/c
ID   ABL06815 standard; cDNA; 1788 BP.
XX
AC   ABL06815;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 14927.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB62712.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 14927; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1788 BP; 539 A; 579 C; 432 G; 238 T; 0 U; 0 Other;

  Query Match             1.1%;  Score 43;  DB 4;  Length 1788;
  Best Local Similarity   47.3%;  Pred. No. 0.67;
  Matches  130;  Conservative    0;  Mismatches  145;  Indels    0;  Gaps    0;

Qy         58 GCTGATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTTGTGGTTTGTGGTTTGTG 117
              ||||  |    |   | ||||        ||| ||  | | ||  ||||    |  |  |
Db        646 GCTGTTGTTGTTGTTGCTGTTGCTGGCGCAGCGCCTGCGGTTGTTGGTAGTGTGCATAGG 587

Qy        118 GTTTGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGA 177
              | | ||||      || || |||||    |   |  |||  |    |  |    |  | |
Db        586 GATCGTGCAGTTGCTGCTGCTGTGCCCGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGAA 527

Qy        178 TATTGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATC 237
              |   | | ||      | | ||   |||||||  |  | ||    | ||   ||||
Db        526 GTTGTTGGTATGTGGGCGTTATGGCCGCTGTGTTGTATTGCTGGTGTTGTATGGGTGGCG 467

Qy        238 GATATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATG 297
              |   |||  |||||| |  | |   ||  | | | |||||| ||   |      |
Db        466 GTGGCAGAGTCTGTTGCTGCTGTTGCAGATAGTGTTGGTGTTGGTGTTGCTGCTGCTGCA 407

Qy        298 CTTGCTCGTGGTGGAACTGGTGCTGTAGCGCCAGT 332
              |||||  ||  ||   ||||| |||  ||  | ||
Db        406 GTTGCTGCTGATGTTGCTGGTACTGCTGCTGCTGT 372


RESULT 11
ABL06814/c
ID   ABL06814 standard; cDNA; 4773 BP.
XX
AC   ABL06814;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 14924.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
```

XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB62711.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 14924; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4773 BP; 1400 A; 1213 C; 989 G; 1171 T; 0 U; 0 Other;

Query Match 1.1%; Score 43; DB 4; Length 4773;
Best Local Similarity 47.3%; Pred. No. 1.2;
Matches 130; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```
Qy        58 GCTGATTGATTGTCTTCTGTGATATATCCAGAGCTCGTGTTTTGTGGTTTGTGGTTTGTG 117
             ||||  |    |   | ||||        ||| ||  | | ||  ||||     |  |  |
Db      2965 GCTGTTGTTGTTGTTGCTGTTGCTGGCGCAGCGCCTGCGGTTGTTGGTAGTGTGCATAGG 2906

Qy       118 GTTTGTGCTTGGATTGTTGATGTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGA 177
             | | ||||      || || |||||    |   |  ||| |     |  |    | |  |
Db      2905 GATCGTGCAGTTGCTGCTGCTGTGCCCGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGAA 2846

Qy       178 TATTGAGTTGTGCCTCGGCTGTGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATC 237
               |  | ||      | | ||   |||||||  |  | ||    | ||   ||||
Db      2845 GTTGTTGGTATGTGGGCGTTATGGCCGCTGTGTTGTATTGCTGGTGTTGTATGGGTGGCG 2786

Qy       238 GATATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATG 297
             |    |||  |||||| |  | |   ||  | | | |||||| ||   |      |
Db      2785 GTGGCAGAGTCTGTTGCTGCTGTTGCAGATAGTGTTGGTGTTGGTGTTGCTGCTGCTGCA 2726

Qy       298 CTTGCTCGTGGTGGAACTGGTGCTGTAGCGCCAGT 332
              |||||  || ||   ||||| |||  ||  | ||
Db      2725 GTTGCTGCTGATGTTGCTGGTACTGCTGCTGCTGT 2691
```

RESULT 12
AAC44195
ID AAC44195 standard; DNA; 574 BP.
XX
AC AAC44195;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41975.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 1.1%; Score 42.8; DB 3; Length 574;
Best Local Similarity 57.5%; Pred. No. 0.38;
Matches 77; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy       1251 GACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGA 1310
              || |   |||||||||||||||||| || || | |  ||||        ||   | | |||
Db        122 GAAGTTGAAGAAGTTGATGAAGAAGAAGAAGAAAAAAGAGGAAATGTGATGATGTTGA 181

Qy       1311 TGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAAT 1370
              || |   ||||       ||  ||||| |||    || ||| ||  | ||  | ||  | |
Db        182 TGATGATGATGGTGGAATTGCTCGGAAAAAAAGCAAGAAGATGAAGATTTCAACATCAGT 241

Qy       1371 TATAAATGCTAACC 1384
              |  | |   ||| |
Db        242 TTCAGACAATAATC 255
```

RESULT 13
ABQ69245_18
Continuation (19 of 31) of ABQ69245 from base 1800001 (Listeria innocua DNA sequence #6
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245

	Fragment Name	Begin	End
WP	ABQ69245_00	1	110000
WP	ABQ69245_01	100001	210000
WP	ABQ69245_02	200001	310000
WP	ABQ69245_03	300001	410000

```
WP      ABQ69245_04         400001    510000
WP      ABQ69245_05         500001    610000
WP      ABQ69245_06         600001    710000
WP      ABQ69245_07         700001    810000
WP      ABQ69245_08         800001    910000
WP      ABQ69245_09         900001   1010000
WP      ABQ69245_10        1000001   1110000
WP      ABQ69245_11        1100001   1210000
WP      ABQ69245_12        1200001   1310000
WP      ABQ69245_13        1300001   1410000
WP      ABQ69245_14        1400001   1510000
WP      ABQ69245_15        1500001   1610000
WP      ABQ69245_16        1600001   1710000
WP      ABQ69245_17        1700001   1810000
WP      ABQ69245_18        1800001   1910000
WP      ABQ69245_19        1900001   2010000
WP      ABQ69245_20        2000001   2110000
WP      ABQ69245_21        2100001   2210000
WP      ABQ69245_22        2200001   2310000
WP      ABQ69245_23        2300001   2410000
WP      ABQ69245_24        2400001   2510000
WP      ABQ69245_25        2500001   2610000
WP      ABQ69245_26        2600001   2710000
WP      ABQ69245_27        2700001   2810000
WP      ABQ69245_28        2800001   2910000
WP      ABQ69245_29        2900001   3010000
WP      ABQ69245_30        3000001   3011208

  Query Match             1.1%;  Score 42.2;  DB 6;  Length 110000;
  Best Local Similarity   47.0%;  Pred. No. 14;
  Matches  131;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy        814 AAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGATTTGCAAGGGCCAGCC 873
              || ||   ||    |||    |   |  ||   ||||  ||||  |     |    ||
Db      27998 AACAAAAAAGGACATGAATTTATGGCCAATATGCCAACAGAAGAAGTTTTCTGTTGTGCG 28057

Qy        874 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 933
              | ||    | ||||  ||   ||||  ||     ||  |   |  | |||  ||   |
Db      28058 GATAAACTAAAAGTGGAAGGTTATGTTTCAAGCACGAAACCACTTAGTTATGCAGGAAAT 28117

Qy        934 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 993
              |   |||  |    | | ||  ||  |   | | ||      | ||  ||  | | || || ||
Db      28118 ATTATTGATGACTTCAAAATCACTTTCGAAAAAGGTCGTATTGTTGGCGTTGAAGCTGCA 28177

Qy        994 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 1053
              ||      | || |||   |||   |  ||  ||  ||||| |    ||     |
Db      28178 TCTGGCGAAGAAATTTTAAAAGATTTAATCGCAACGGATGAAGGTTCTCATTATTTAGGC 28237

Qy       1054 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAA 1092
              |||| | |||| ||| | || |  | |  | |  || ||
Db      28238 GAAGTGGCTCTAGTTCCAGATCCATCCCCTATTTCCCAA 28276


RESULT 14
ABQ67194/c
ID   ABQ67194 standard; DNA; 319630 BP.
XX
AC   ABQ67194;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria innocua contig DNA sequence #7.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria innocua.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR003061.
XX
PR   04-OCT-2000; 2000FR-00012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators.
XX
PS   Claim 5; SEQ ID NO 7; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences (ABQ67188-
CC   ABQ71212) from Listeria sp. The sequences are useful as probes and
CC   primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of gene
CC   expression. Proteins encoded by the nucleic acid sequences can be used to
CC   screen for compounds that modulate gene expression, replication and
CC   pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in anti-
CC   Listeria vaccines. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 319630 BP; 105207 A; 55428 C; 66726 G; 92263 T; 0 U; 6 Other;

  Query Match             1.1%;  Score 42.2;  DB 6;  Length 319630;
  Best Local Similarity   47.0%;  Pred. No. 28;
  Matches  131;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy        814 AAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGATTTGCAAGGGCCAGCC 873
              || ||   ||    |||    |   |  ||   ||||  ||||  |     |    ||
Db     289425 AACAAAAAAGGACATGAATTTATGGCCAATATGCCAACAGAAGAAGTTTTCTGTTGTGCG 289366

Qy        874 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 933
              | ||    | ||||  ||   ||||  ||     ||  |   |  | |||  ||   |
Db     289365 GATAAACTAAAAGTGGAAGGTTATGTTTCAAGCACGAAACCACTTAGTTATGCAGGAAAT 289306

Qy        934 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 993
              |   |||  |    | | ||  ||  |   | | ||      | ||  ||  | | || || ||
Db     289305 ATTATTGATGACTTCAAAATCACTTTCGAAAAAGGTCGTATTGTTGGCGTTGAAGCTGCA 289246

Qy        994 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 1053
              ||      | || |||   |||   |  ||  ||  ||||| |    ||     |
Db     289245 TCTGGCGAAGAAATTTTAAAAGATTTAATCGCAACGGATGAAGGTTCTCATTATTTAGGC 289186

Qy       1054 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAA 1092
              |||| | |||| ||| | || |  | |  | |  || ||
Db     289185 GAAGTGGCTCTAGTTCCAGATCCATCCCCTATTTCCCAA 289147


RESULT 15
AAS90679
ID   AAS90679 standard; cDNA; 246 BP.
XX
AC   AAS90679;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #26483.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
```

XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG26492.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 26483; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 246 BP; 126 A; 20 C; 77 G; 23 T; 0 U; 0 Other;

Query Match 1.1%; Score 41.4; DB 5; Length 246;
Best Local Similarity 60.0%; Pred. No. 0.58;
Matches 69; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
Qy       1178 AGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCA 1237
             || | |||||  |      | ||  ||   || |||||||  |     | |||||     |
Db         71 AGGACTGTGAAGAAGAAAGAAAGAAGGACTGCGAAGAAGAAGAAGAAAGAAGAAAGAAGA 130

Qy       1238 AATCTAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAA 1292
              |    |||| ||||| | | ||||||  || |||||||  || || | |  |||||
Db        131 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 185
```

Search completed: March 7, 2006, 17:08:55
Job time : 2159 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 7, 2006, 12:35:33 ; Search time 14346 Seconds
(without alignments)
12706.153 Million cell updates/sec

Title: US-09-828-068-1
Perfect score: 3896
Sequence: 1 cgcggctgtcggagcaaacg.....gtcaacaccggagaatttac 3896

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

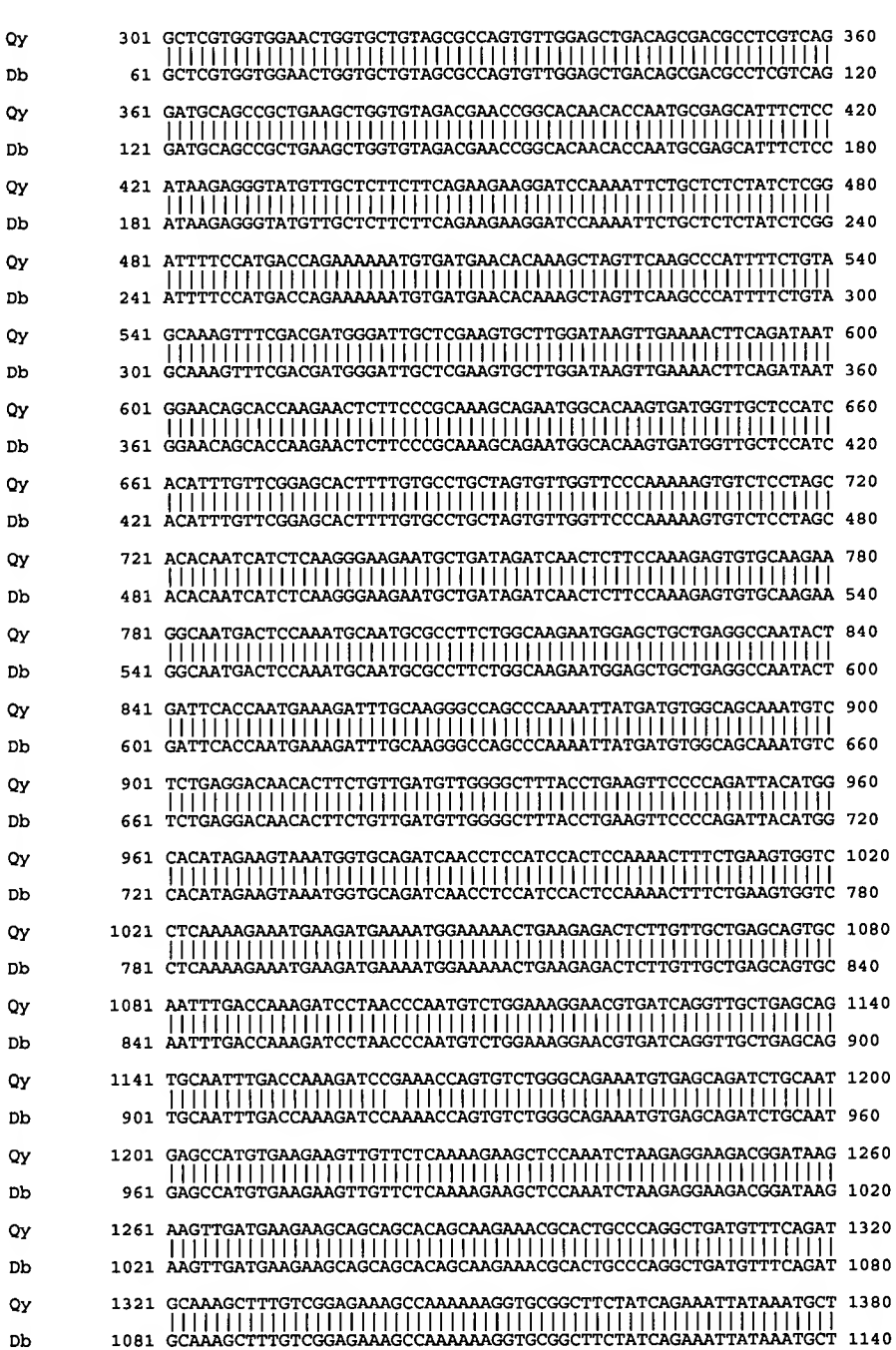
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3159.6	81.1	3174	10	CL958119	CL958119 OsIFCC000
c	2	828	21.3	884	10	AG906393	AG906393 Oryza sat
	3	635.2	16.3	838	10	CW621556	CW621556 OP__Ba001
c	4	528.6	13.6	904	10	CZ880884	CZ880884 OC__Ba028
c	5	524.2	13.5	723	10	CW759415	CW759415 OG_BBa006
	6	473	12.1	857	10	CZ694942	CZ694942 OC__Ba000
	7	457.4	11.7	687	10	CL602016	CL602016 OB__Ba007
	8	454.2	11.7	739	10	CZ785763	CZ785763 OC__Ba014
	9	387.2	9.9	460	10	CZ187525	CZ187525 OA_BBa011
c	10	361.8	9.3	384	10	CW686090	CW686090 OG_BBa004
c	11	355.6	9.1	1025	9	BZ795917	BZ795917 PUFHM45TD
	12	352	9.0	888	9	CC355521	CC355521 PUHPL93TB
c	13	327	8.4	731	10	CL704355	CL704355 SP__Bb001
	14	315.6	8.1	647	10	CL597614	CL597614 OB__Ba005
	15	268.6	6.9	491	10	AG209720	AG209720 Oryza sat
	16	268.2	6.9	523	3	BJ246359	BJ246359 BJ246359
c	17	243.8	6.3	587	10	CL703548	CL703548 SP__Bb000
	18	237.4	6.1	885	10	CG091207	CG091207 PUJFK24TD
	19	236	6.1	469	10	AG213697	AG213697 Oryza sat
	20	229.6	5.9	593	3	BQ244810	BQ244810 TaE15035B
c	21	217.4	5.6	389	10	CZ821320	CZ821320 OC__Ba019
	22	205.8	5.3	999	9	BZ795914	BZ795914 PUFHM45TB
	23	205.2	5.3	556	5	BU974305	BU974305 HB27I10r
	24	205	5.3	787	10	CG345663	CG345663 OGWLW77TH
	25	178.6	4.6	469	1	AL818058	AL818058 AL818058
	26	173.6	4.5	641	9	BZ411511	BZ411511 OGACC43TM
c	27	172	4.4	849	9	BZ411504	BZ411504 OGACC43TC
c	28	158.6	4.1	674	10	CW161591	CW161591 104_568_1
	29	153.2	3.9	718	10	CW161592	CW161592 104_568_1
c	30	153.2	3.9	736	10	CW365639	CW365639 fsbb001f0
c	31	152	3.9	766	9	CC603916	CC603916 OGWHM55TV
c	32	139.4	3.6	635	9	BH880296	BH880296 ht52d11.b
	33	129	3.3	822	10	CG365016	CG365016 OGHAD70TV
	34	128	3.3	668	10	CW419078	CW419078 fsbb001f1
c	35	124.4	3.2	753	9	CC160109	CC160109 ig06g11.b
	36	124.4	3.2	995	10	CG451574	CG451574 OG8AD34TV
	37	123.8	3.2	276	10	CZ261039	CZ261039 OA_BBa010
c	38	122.8	3.2	842	9	BZ705032	BZ705032 PUCCS90TD
	39	120.8	3.1	703	9	BZ776688	BZ776688 ii04h10.b
c	40	119.6	3.1	688	9	BZ776689	BZ776689 ii04h10.g
c	41	118.8	3.0	970	10	CG032005	CG032005 PUFTA25TD
c	42	114.2	2.9	653	9	BZ311518	BZ311518 ic68h08.b
c	43	114.2	2.9	870	10	CG091452	CG091452 PUFPF81TB
c	44	111.2	2.9	903	9	BZ704364	BZ704364 PUBLF78TD
c	45	109	2.8	664	3	BQ242025	BQ242025 TaE15035B

ALIGNMENTS

RESULT 1
CL958119
LOCUS CL958119 3174 bp DNA linear GSS 21-SEP-2004
DEFINITION OsIFCC000847 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL958119
VERSION CL958119.1 GI:52370980
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 3174)
AUTHORS Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES Location/Qualifiers
source 1. .3174
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Query Match 81.1%; Score 3159.6; DB 10; Length 3174;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3165; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 60


Qy 301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360
Db 61 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 120
Qy 361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420
Db 121 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 180
Qy 421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480
Db 181 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 240
Qy 481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540
Db 241 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 300
Qy 541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600
Db 301 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 360
Qy 601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660
Db 361 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 420
Qy 661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720
Db 421 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 480
Qy 721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780
Db 481 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 540
Qy 781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840
Db 541 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 600
Qy 841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900
Db 601 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 660
Qy 901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960
Db 661 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 720
Qy 961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020
Db 721 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 780
Qy 1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080
Db 781 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 840
Qy 1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140
Db 841 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 900
Qy 1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200
Db 901 TGCAATTTGACCAAAGATCCAAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 960
Qy 1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260
Db 961 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1020
Qy 1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320
Db 1021 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1080
Qy 1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380
Db 1081 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1140
Qy 1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440
Db 1141 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1200
Qy 1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500
Db 1201 GAGGATGATATAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTATC 1260
Qy 1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560
Db 1261 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAAAACAAGACAAAACGCAAATAC 1320
Qy 1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620
Db 1321 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1380
Qy 1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
Db 1381 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1440
Qy 1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
Db 1441 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1500
Qy 1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
Db 1501 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1560
Qy 1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
Db 1561 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1620
Qy 1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
Db 1621 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1680
Qy 1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
Db 1681 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCGCTCGGCAAAGGTTTCTCCAGCTGAG 1740
Qy 1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
Db 1741 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 1800
Qy 2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
Db 1801 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 1860
Qy 2101 GTTGAACTGCTAGCTAAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
Db 1861 GTTGAACTGCTAGCTAAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 1920
Qy 2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
Db 1921 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 1980
Qy 2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
Db 1981 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2040
Qy 2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
Db 2041 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2100
Qy 2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
Db 2101 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2160
Qy 2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
Db 2161 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2220
Qy 2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520

```
Db      2221 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2280

Qy      2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
Db      2281 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2340

Qy      2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
Db      2341 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2400

Qy      2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
Db      2401 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCTGTTAACCAGGTCCATGCA 2460

Qy      2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760
Db      2461 AGAAATTTTCCAAGCACAATAGCAACCATGGAAACGAGTAAGTTATGTGATCGGAGAAAT 2520

Qy      2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
Db      2521 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2580

Qy      2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
Db      2581 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2640

Qy      2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
Db      2641 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2700

Qy      2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
Db      2701 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCGCGGCAT 2760

Qy      3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
Db      2761 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGATTTGGTGTGCTTGGCTCC 2820

Qy      3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
Db      2821 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 2880

Qy      3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
Db      2881 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 2940

Qy      3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
Db      2941 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3000

Qy      3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
Db      3001 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3060

Qy      3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
Db      3061 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3120

Qy      3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGA 3414
Db      3121 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGA 3174


RESULT 2
AG906393/c
LOCUS       AG906393               884 bp    DNA     linear   GSS 03-NOV-2004
DEFINITION  Oryza sativa (indica cultivar-group) genomic DNA, BAC end sequence,
            BAC clone:K0471H01_R, genomic survey sequence.
ACCESSION   AG906393
VERSION     AG906393.1  GI:55372775
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Katagiri,S., Wu,J., Ito,Y., Karasawa,W., Shibata,M., Kanamori,H.,
            Katayose,Y., Namiki,N., Matsumoto,T. and Sasaki,T.
  TITLE     End Sequencing and Chromosomal in silico Mapping of BAC Clones
            Derived from an indica Rice Cultivar, Kasalath
  JOURNAL   Breeding Science 54, 273-279 (2004)
REFERENCE   2  (bases 1 to 884)
  AUTHORS   Sasaki,T., Matsumoto,T. and Wu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-2004) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki, 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The orientation of the sequence is from SP6 side of the BAC clone.
FEATURES             Location/Qualifiers
     source          1. .884
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Kasalath"
                     /db_xref="taxon:39946"
                     /clone="K0471H01_R"
ORIGIN

  Query Match             21.3%;  Score 828;  DB 10;  Length 884;
  Best Local Similarity   96.8%;  Pred. No. 1.5e-220;
  Matches  854;  Conservative    0;  Mismatches   27;  Indels    1;  Gaps    1;

Qy      2033 AAAAGAAGCAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCA 2092
Db       881 AGAAAAAGACGCAAAACTGAAGGTGTCGGTNAAACCAGGCCATGATAGATG-CCTTCCCA 823

Qy      2093 TGGATATTGTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTG 2152
Db       822 TGGATATTGTTGAACTGCAAGCTAAAACCCAGCATGAGAGGCAGCTTATGACTGAGACTG 763

Qy      2153 ATTGTTCTGACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAG 2212
Db       762 ATTGTTCTGACATCAACCGTATTCAATCCAAGCCAACTGCTGATGATGATTGTGTNATAG 703

Qy      2213 TAGCTGCCAAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGA 2272
Db       702 TAGCTGCCAAGGATGGTTCAGATAATGCATCAAGTGTGTTTGACACTAATTCCCAACAGA 643

Qy      2273 AGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACAC 2332
Db       642 AGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACAC 583

Qy      2333 AAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGA 2392
Db       582 AAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGA 523

Qy      2393 TGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGT 2452
Db       522 TGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGT 463

Qy      2453 ATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGG 2512
Db       462 ATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGG 403

Qy      2513 AGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGAC 2572
Db       402 AGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGAC 343

Qy      2573 CTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATG 2632
Db       342 CTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATG 283

Qy      2633 CATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGG 2692
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGG 223

Qy       2693 TCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATC 2752
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 TCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATC 163

Qy       2753 GGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTC 2812
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTC 103

Qy       2813 TGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGA 2872
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 TGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGA 43

Qy       2873 ACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATA 2914
                ||||||||||||||||||||||||||||||||||||||||||
Db         42 ACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACATA 1
```

```
RESULT 3
CW621556
LOCUS       CW621556                838 bp    DNA     linear   GSS 22-OCT-2004
DEFINITION  OP__Ba0019J18.f OP__Ba Oryza punctata genomic clone OP__Ba0019J18
            5', genomic survey sequence.
ACCESSION   CW621556
VERSION     CW621556.1  GI:54547814
KEYWORDS    GSS.
SOURCE      Oryza punctata
  ORGANISM  Oryza punctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   SanMiguel,P., Westerman,R., Kim,H., Yu,Y., Wissotski,M., Yost,D.,
            Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C.,
            Hatfield,J., Soderlund,C., Wing,R. and Jackson,S.A.
  TITLE     OMAP Project - Purdue University
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Scott A. Jackson
            Jackson Laboratory
            Purdue University
            915 W. State St., West Lafayette, IN 47907, USA
            Tel: 7654963621
            Fax: 7654967255
            Email: sjackson@purdue.edu
            Basecalling by phred version 0.020425.c. This sequence was derived
            from the raw sequence read by clipping with lucy version 1.19s.
            Bases 46-883 of the raw sequence (length 1374) were retained after
            clipping.
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 161   Std Error: 0.00
            Plate: 0019  row: J  column: 18
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Oryza punctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4537"
                     /clone="OP__Ba0019J18"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OP__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             16.3%;  Score 635.2;  DB 10;  Length 838;
  Best Local Similarity   94.4%;  Pred. No. 2.3e-166;
  Matches  670;  Conservative    0;  Mismatches   38;  Indels    2;  Gaps    1;
```

```
Qy       2847 CCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTA 2906
                ||||||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||||
Db          1 CCCCAACTATGTAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTTTCAGTA 60

Qy       2907 TGCACATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAATGGAAA 2966
                ||||| |||||||||||||||||||| |||||||||||||||||||| | ||||||||
Db         61 TGCACCTAATCAGTACAAAGGATCAACTAGCACATCATATGGCAGTAACATAAATGGAAA 120

Qy       2967 GATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACAGACCTTT 3026
                |||||||||||||||||||||||||||||||||||| |||||||| ||||||||||||
Db        121 GATTCCATTGACATTCGAAGACTTATCACGGCATCAGATGCATGATTTGCACAGACCTTT 180

Qy       3027 ACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTG 3086
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTG 240

Qy       3087 GTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATC 3146
                |||||||||||||||||||||||||||||||||||| |||||||| ||||| ||||||||
Db        241 GTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTGGGAGTGTCGACAGGGATAACATC 300

Qy       3147 GCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAA 3206
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        301 GCATCAGATGAACAGAAAGGAACATTTTGAAGCCTTGAATTCTGGAATGTTTTCAGCAAA 360

Qy       3207 ATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAA 3266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAA 420

Qy       3267 CAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATCGGTT 3326
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        421 CAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATAGGTT 480

Qy       3327 TGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATTTTACTACAATCAGTAA 3386
                |||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||||
Db        481 TGTGAGACAGGATATCTGTGTAACTAACAAGAATCCAGCTGATTTTACTACAATCAGTAA 540

Qy       3387 CGATAACGAGTATATGGATTACCGCTGAAGCAGAAAGTGGTGTGCATAATTCCTGAACAT 3446
                |||||||||||||||||||||||||||| |||||||||||||||||||||||||    |||||
Db        541 CGATAACGAGTATATGGATTACCGGTGAAGCAGAAAGTGGTGTGCATAATTCTGAAACAT 600

Qy       3447 TTACAATCATACATTTCATCTTTATGGCGCCAAATAGTCATACTGTA--AGAGGAGGGCT 3504
                |||||||||||||||||||||||||||||||||| ||||||||||||  ||||||||||||
Db        601 TTACAATCATACATTTCATCTTTATGGCGCCAAACAGTCATACTGTAAGAGAGGAGGGCT 660

Qy       3505 TTGCTGGATCTGCTGTAAGGCTTCTTGTAAGTTGTGGATGCCCCATTTTC 3554
                |||||||||||||||||||    |||    || |   |    ||  |||
Db        661 TTGCTGGATCTGCTGTAAGGTAAGTTGATTTTTTTTCTTCTTGCAAGTTC 710
```

```
RESULT 4
CZ880884/c
LOCUS       CZ880884                904 bp    DNA     linear   GSS 27-JUL-2005
DEFINITION  OC__Ba0283G18.f OC__Ba Oryza coarctata genomic clone OC__Ba0283G18
            5', genomic survey sequence.
ACCESSION   CZ880884
VERSION     CZ880884.1  GI:71356640
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
```

```
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0283  row: G  column: 18
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0283G18"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             13.6%;  Score 528.6;  DB 10;  Length 904;
  Best Local Similarity   77.6%;  Pred. No. 2.2e-136;
  Matches  750;  Conservative    0;  Mismatches   64;  Indels  153;  Gaps    3;

Qy        447 GAAGAAGGATCCAAAATTCTGCTCTCTATCTCGGATTTTCCATGACCAGAAAAAATGTGA 506
              |||||||||||||||| | |||||||||||||||||||||| ||||||||||||||||||
Db        904 GAAGAAGGATCCAAAACTTTGCTCTCTATCTCGGATTTTCCGTGACCAGAAAAAATGTGA 845

Qy        507 TGAACACAAAGCTAGTTCAAGCCCATTTTCTGTAGCAAAGTTTCGACGATGGGATTGCTC 566
              ||| ||| ||| |||||||||||||||| |||||||||||||||| ||||||||||
Db        844 TGATCACCAAGTTAGTTCAAGCCCATTTTTTGTAGCAAAGTTTCGACGGTGGGATTGCTC 785

Qy        567 GAAGTGCTTGGATAAGTTGAAAACTTCAGATAATGGAACAGCACCAAGAACTCTTCCCGC 626
              |||||||||||||||||||||||||||||| |||||||||| |||||||| ||| ||||
Db        784 TAAGTGCTTGGATAAGTTGAAAACTTCAGATCATGGAACAGCTCCAAGAACACTTTCCGC 725

Qy        627 AAAGCAGAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGCC 686
              ||||||| | ||||||||||||||||||||||||||||||||||||||||||| ||||
Db        724 AAAGCAGGACGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTTTGCC 665

Qy        687 TGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAATGC 746
              |||| |||||||||| |||  |||||||||||||||||||||||||||||||||| |||
Db        664 GGCTACTGTTGGTTCCAAAAGTGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAGTGC 605

Qy        747 TGATAGATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGCC 806
              |||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||
Db        604 TGATAGATCAACTCTTCCGAAGAGTGCGCAAGAAGGCAATGACTCCAAATGCAATGCGCC 545

Qy        807 TTCTGGCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA------------ 854
              |||||||||||||||||||||||||||||| |||||||||||| |||||
Db        544 TTCTGGCAAGAATGGAGCTGCTGAGGCCAGTACTGATTCACCCATGAAAGGTATGGTAGA 485

Qy        855 ------------------------------------------------------------ 854

Db        484 TGTAGAGCCTTTTAAATTCCTAGTAAGATTTTATTTAAGGATAGAATAAGATGATGTTTG 425

Qy        855 ------------AGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTC 902
                         ||||||||| |||||||||||||||||||| ||||||||||| ||||
Db        424 TGTGATTTTTTCAGATTTGCATGGGCCAGCCCAAAATTATGATATGGCAGCAAATATCTC 365

Qy        903 TGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCA 962
              ||| |||||||||||||| |||||||||||||||||||||||||||||||||| ||||||
Db        364 CGAGAACAACACTTCTGTAGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTATATGGCA 305

Qy        963 CATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCT 1022
              || | ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        304 CAAAAAAGGAAATGGTGCCGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCT 245

Qy       1023 CAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAA 1082
              |||| |||||||||||||||||||||||||||||||| ||||
Db        244 CAAACGAAATGAAGATGAAAATGGAAAAACTGAAGAGATTCTT---------------- 202

Qy       1083 TTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAGTG 1142
                                                              ||||||||||||||
Db        201 ----------------------------------------------ATTGCTGAGCAGTG 188

Qy       1143 CAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAATGA 1202
              |||||| ||||| ||||||||||||| ||||||| |||||||||| ||| |||||||
Db        187 CAATTTTACCAATGATCCGAAACCAATGTCTGGACAGAAATGTGATCAGGTCTGCAACAG 128

Qy       1203 GCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATC------TAAGAGGAAGACGGA 1256
                    |||||||||| |||||||||||||||||      ||||||||||||||||
Db        127 TGAGCCATGTGAAGTTGTTCTGAAAAGAAGCTCCAAATCTAAATGTAAGAGGAAGACGGA 68

Qy       1257 TAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTC 1316
              |||||||||||||||||||| ||||||| ||||||||||||||||||||||||||| |||
Db         67 TAAGAAGTTGATGAAGAAGAAGCAGCAAAGCAAGAAACGCACTGCCCAGGCTGATGATTC 8

Qy       1317 AGATGCA 1323
              ||||| |
Db          7 AGATGTA 1


RESULT 5
CW759415/c
LOCUS       CW759415                723 bp    DNA     linear   GSS 09-NOV-2004
DEFINITION  OG_BBa0067J13.f OG_BBa Oryza glaberrima genomic clone OG_BBa0067J13
            5', genomic survey sequence.
ACCESSION   CW759415
VERSION     CW759415.1  GI:55603764
KEYWORDS    GSS.
SOURCE      Oryza glaberrima (African rice)
  ORGANISM  Oryza glaberrima
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 723)
  AUTHORS   Kim,H., Yu,Y., Wissotski,M., Byrne,M., Stum,D., Smart,D., Rao,K.,
            Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C.
            and Wing,R.
  TITLE     OMAP
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0067  row: J  column: 13
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .723
                     /organism="Oryza glaberrima"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4538"
                     /clone="OG_BBa0067J13"
                     /tissue_type="young leaves"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OG_BBa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             13.5%;  Score 524.2;  DB 10;  Length 723;
  Best Local Similarity   87.6%;  Pred. No. 3.5e-135;
  Matches  633;  Conservative    0;  Mismatches    3;  Indels   87;  Gaps    2;

Qy        633 GAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTGTGCCTGCTAG 692
              |||||||||||||||||||||||||||||||||||||||||| |||||||||  ||||||||
Db        723 GAATGGCACAAGTGATGGTTGCTCCATCACATTTGTTTGGAGCACTTTTGGCCCTGCTAG 664
```

```
Qy        693 TGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAG-AATGCTGATA 751
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        663 TGTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAAATGCTGATA 604

Qy        752 GATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGCCTTCTG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGCCTTCTG 544

Qy        812 GCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA---------------- 854
              |||||||||||||||||||||||||||||||||||||||||
Db        543 GCAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAAAGGTATGGTAGATGTAG 484

Qy        855 ------------------------------------------------------------ 854

Db        483 AGCCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGAATAAACTAATGTTTGTGT 424

Qy        855 ---------AGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTCTGA 905
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTCTGA 364

Qy        906 GGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCACAT 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCACAT 304

Qy        966 AGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCTCAA 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 AGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCTCAA 244

Qy       1026 AAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAATTT 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAATTT 184

Qy       1086 GACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAGTGCAA 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAGTGCAA 124

Qy       1146 TTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAATGAGCC 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAATGAGCC 64

Qy       1206 ATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAGAAGTT 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 ATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAGAAGTT 4

Qy       1266 GAT 1268
              |||
Db          3 GAT 1
```

RESULT 6
CZ694942

```
LOCUS       CZ694942                857 bp    DNA     linear   GSS 25-JUL-2005
DEFINITION  OC__Ba0007O13.f OC__Ba Oryza coarctata genomic clone OC__Ba0007O13
            5', genomic survey sequence.
ACCESSION   CZ694942
VERSION     CZ694942.1  GI:71094090
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 857)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0007  row: O  column: 13
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .857
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0007O13"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
```

Query Match 12.1%; Score 473; DB 10; Length 857;
Best Local Similarity 85.5%; Pred. No. 9.2e-121;
Matches 553; Conservative 0; Mismatches 85; Indels 9; Gaps 2;

```
Qy       1324 AAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCTAAC 1383
              | ||||||| |||||||||| |||||||| |||||| |||| |||||||||||| ||| ||
Db          2 ATGCTTTGTTGGAGAAAGCCGAAAAAGGTACGGCTTTTATCGGAAATTATAAAAGCTGAC 61

Qy       1384 CAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGTGAG 1443
              || | ||||||| |||||||||| ||||||||| ||||||||  ||||||||||||||||
Db         62 CAAGGTGAGGATCCTAGAAGTGATAAAGTTCATCATGAAAATGTTGCTGATCCCTGTGAG 121

Qy       1444 GATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGCAAC 1503
              ||||| |||||       |    ||||  |||||||||||||||||||||||| |||| ||
Db        122 GATGAGAGAAG------CTTAATCCCATTGGAAGTAAGCATGGATATTCCTGGTAGCCAC 175

Qy       1504 CATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATACTCT 1563
              || | ||||||||| ||||||| |||||||| |||||||||||||| |||||||| | ||
Db        176 CAGAAAGTGGGAGATGATGGGTCAAAATCAACTAAGAACAAGACAGAACGCAAATGCACT 235

Qy       1564 GATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAA---GA 1620
              |||| |||||||||||||||||||| |||||||||||||||||||| |||||||     |
Db        236 GATGCTGTAGATGATGGATCATCACTCATGAACTGGCTGAATGGAAAGAAGAAAAAAGTA 295

Qy       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
              |||| || |||||||||||| | ||  |||||||||||||||||||||||| | |||||||
Db        296 ACTGAAAATGTGCATCACACTGCTGTACATCCAGCTGGGAATTTGAGCAATAGAAAAGTG 355

Qy       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
              || |||| ||  ||| |||||||||||||||| || ||| | ||||||||||||||| |
Db        356 ACCCCCAGTGTCAGTGCTCAGCATGATGATGAGGATAATATTCAAAATGGTCTTGACATA 415

Qy       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
              |||||||||||||| ||| |||| || |||| |||||||| ||| |||||||||| ||
Db        416 AATATGCATAAGACTGATATCTGCCAACATGAATCAGAAAACTCTACACAGAGGTGTTCG 475

Qy       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| |||||
Db        476 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGTGGAAAACACATTCAACTGCCAGTGCCAAA 535

Qy       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
              |||||| ||||| ||||||||  ||||||| |||| | ||| |||  ||||||||||
Db        536 AATGGTGATGAAAACACCAGAAACAGTCAGAATATACCTATACCCAGGACAGAAGATCAA 595

Qy       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGT 1967
              |||||| |||||||||||||||||||||||||   || | |||||||
Db        596 TGCCAGGTGGAAACCGAAAACTCTGTTCTGAGGTGCTTGACAAAGGT 642
```

RESULT 7
CL602016

```
LOCUS       CL602016                687 bp    DNA     linear   GSS 10-JUN-2004
```

```
DEFINITION  OB__Ba0072E10.r OB__Ba Oryza brachyantha genomic clone
            OB__Ba0072E10 3', genomic survey sequence.
ACCESSION   CL602016
VERSION     CL602016.1  GI:48579779
KEYWORDS    GSS.
SOURCE      Oryza brachyantha
  ORGANISM  Oryza brachyantha
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE     OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 145000   Std Error: 0.00
            Plate: 0072  row: E  column: 10
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Oryza brachyantha"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4533"
                     /clone="OB__Ba0072E10"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OB__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match            11.7%;  Score 457.4;  DB 10;  Length 687;
  Best Local Similarity   84.4%;  Pred. No. 2.1e-116;
  Matches  541;  Conservative    0;  Mismatches   91;  Indels    9;  Gaps    2;

Qy       1330 TGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCTAACCAGGTT 1389
             |||||||||||||| |||||||| |||||||||||||||||||||||||||||| | |
Db             1 TGTCGGAGAAAGCCGAAAAAGGTACGGCTTCTATCAGAAATTATAAATGCTAACCAAGGT 60

Qy       1390 GAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGTGAGGATGAT 1449
             | |||  ||||||| || |||| ||||| ||||||||   |||||| | ||||||||||
Db            61 GGGGAACCTAGAAGCGATGAAGATCATCATGAAAATGTTTCTGATCACCGTGAGGATGAG 120

Qy       1450 AGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGCAACCATACA 1509
             |||||    |      ||||  |||||||||||||||||| |||||| |||| |||| | |
Db           121 AGAAGGTTAA------TCCCATTGGAAGTAAGCATGGATTTTCCTGGTAGCCACCAGAAA 174

Qy       1510 GTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATACTCTGATGTT 1569
             ||  |||||||||  ||||| |||| ||||| |||||||||||||||| | |||||| |
Db           175 GTAAGAGAAGATGAATTAAAGTCAACTAAGAGCAAGACAAAACGCAAATTCGCTGATGCT 234

Qy       1570 GTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAG---AAAAGAACTGGA 1626
             |||||||||||||||||| |||||||||||||||||||||||||    |||| ||||| |
Db           235 GTAGATGATGGATCATCACTCATGAACTGGCTGAATGGAAAAAAGAGAAAAACAACTGAA 294

Qy       1627 AGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTGACACCC 1686
             | ||||||||||  ||||  | |||||||||||||||||||||||| |||||||||  |||
Db           295 AATGTGCATCACAGTGTTGTACGTCCAGCTGGGAATTTGAGCAACAGAAAAGTGACCCCC 354

Qy       1687 ACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACAAATATG 1746
             ||||  ||| ||||||||||||||||  || ||| ||||||||||||||||  ||| |||
Db           355 ACTGTCAGTGCTCAGCATGATGATGACGATAATATTGAAAATGGTCTTGACGTAAACATG 414

Qy       1747 CATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCATCAAAG 1806
             |||| ||||||| |  | || |||| |||||| | ||| |||||||||||||| ||||||
Db           415 CATATGACAGATATGCGCCAACATGAATCAGAGAACTCTACACAGAGGTGCTCTTCAAAG 474

Qy       1807 GGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAATATGGT 1866
             || | ||| |||||||||||||| |||||||||||| |||| ||| ||||| |||||
Db           475 GGAACAACGGCGGGTTTGAGTAAGTGGAAAACACATTCACCTGCCAGTGCCAAAAATGGT 534

Qy       1867 GGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAATGCCAG 1926
             | ||||| ||| ||| |  |||||||||||| | |||||||  ||||||||||||||||||
Db           535 GATGAAAACACTAGAGACAGTCAGAACATACCTATACTCAGGACAGAAGATCAATGCCAG 594

Qy       1927 ATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGT 1967
             |||||||||||||||||||||||||||   ||||||||||||
Db           595 ATGGAAACCGAAAACTCTGTTCTGAGGTGCTCGGCAAAGGT 635


RESULT 8
CZ785763
LOCUS       CZ785763               739 bp    DNA     linear   GSS 26-JUL-2005
DEFINITION  OC__Ba0148N23.f OC__Ba Oryza coarctata genomic clone OC__Ba0148N23
            5', genomic survey sequence.
ACCESSION   CZ785763
VERSION     CZ785763.1  GI:71225616
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0148  row: N  column: 23
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0148N23"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match            11.7%;  Score 454.2;  DB 10;  Length 739;
  Best Local Similarity   85.4%;  Pred. No. 1.7e-115;
  Matches  544;  Conservative    0;  Mismatches   83;  Indels   10;  Gaps    3;

Qy       1335 GAGAAAGCCAAAAAAGGTGC-GGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGG 1393
             ||||||||| |||||||| | ||||| |||| |||||||||||| ||| |||| | |||||
Db             1 GAGAAAGCCGAAAAAGGTACAGGCTTTTATCGGAAATTATAAAAGCTGACCAAGGTGAGG 60

Qy       1394 ATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAA 1453
             || |||||||||  |||||||| ||||||||  |||||||||||||||||||| ||||
```

```
Db         61 ATCCTAGAAGTGATAAAGTTCATCATGAAAATGTTGCTGATCCCTGTGAGGATGAGAGAA 120

Qy       1454 GTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGG 1513
              |     |    ||||  ||||||||||||||||||||||||| |||| |||| | |||||
Db        121 G------CTTAATCCCATTGGAAGTAAGCATGGATATTCCTGGTAGCCACCAGAAAGTGG 174

Qy       1514 GAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAG 1573
              |||| ||||||| |||||||| |||||||||||||| ||||||||| | |||||| |||||
Db        175 GAGATGATGGGTCAAAATCAACTAAGAACAAGACAGAACGCAAATGCACTGATGCTGTAG 234

Qy       1574 ATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAA---GAACTGGAAGTG 1630
              |||||||||||||||| |||||||||||||||||||| |||||||     ||||| || ||
Db        235 ATGATGGATCATCACTCATGAACTGGCTGAATGGAAAGAAGAAAAAAGTAACTGAAAATG 294

Qy       1631 TGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTG 1690
              |||||||||| | ||  |||||||||||||||||||||||| | ||||||||| |||| ||
Db        295 TGCATCACACTGCTGTACATCCAGCTGGGAATTTGAGCAATAGAAAAGTGACCCCCAGTG 354

Qy       1691 CGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACAAATATGCATA 1750
              |||  ||||||||||||||||| || ||| | ||||||||||||||| ||||||||||
Db        355 TCAGTGCTCAGCATGATGATGAGGATAATATTCAAAATGGTCTTGACATAAATATGCATA 414

Qy       1751 AGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCATCAAAGGGGA 1810
              ||||  ||| |||| || |||| |||||||| ||| |||||||||| || ||||||||||
Db        415 AGACTGATATCTGCCAACATGAATCAGAAAACTCTACACAGAGGTGTTCGTCAAAGGGGA 474

Qy       1811 AAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAATATGGTGGTG 1870
              |||||||||||||||||||| |||||||||||||| |||| ||| ||||| |||||| ||
Db        475 AAACAGCGGGTTTGAGTAAGTGGAAAACACATTCAACTGCCAGTGCCAAAAATGGTGATG 534

Qy       1871 AAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAATGCCAGATGG 1930
              ||| |||||||||  ||||||| |||| | ||| |||  |||||||||||||||||| |||
Db        535 AAAACACCAGAAACAGTCAGAATATACCTATACCCAGGACAGAAGATCAATGCCAGGTGG 594

Qy       1931 AAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGT 1967
              ||||||||||||||||||||    || | |||||||
Db        595 AAACCGAAAACTCTGTTCTGAGGTGCTTGACAAAGGT 631


RESULT 9
CZ187525
LOCUS       CZ187525                460 bp    DNA     linear   GSS 03-FEB-2005
DEFINITION  OA_BBa0113K02.r OA_BBa Oryza alta genomic clone OA_BBa0113K02 3',
            genomic survey sequence.
ACCESSION   CZ187525
VERSION     CZ187525.1  GI:58544739
KEYWORDS    GSS.
SOURCE      Oryza alta
  ORGANISM  Oryza alta
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   SanMiguel,P., Westerman,R., Kim,H., Yu,Y., Wissotski,M., Yost,D.,
            Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C.,
            Hatfield,J., Soderlund,C., Wing,R. and Jackson,S.A.
  TITLE     OMAP (Oryza Map Alignment Project)- Purdue University
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Scott A. Jackson
            Jackson Laboratory
            Purdue University
            915 W. State St., West Lafayette, IN 47907, USA
            Tel: 7654963621
            Fax: 7654967255
            Email: sjackson@purdue.edu
            Basecalling by phred version 0.020425.c. This sequence was derived
            from the raw sequence read by clipping with lucy version 1.19s.
            Bases 96-555 of the raw sequence (length 1027) were retained after
            clipping.
            Plate: 0113  row: K  column: 02
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Oryza alta"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:52545"
                     /clone="OA_BBa0113K02"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OA_BBa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             9.9%;  Score 387.2;  DB 10;  Length 460;
  Best Local Similarity   94.0%;  Pred. No. 9.6e-97;
  Matches  438;  Conservative    0;  Mismatches   18;  Indels   10;  Gaps    3;

Qy       1330 TGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCTAACCAGGTT 1389
              ||||||||||||||||||||||| |||||||||||||||||||||||  ||||||||||
Db          1 TGTCGGAGAAAGCCAAAAAAGGTCCGGCTTCTATCAGAAATTATAAACACTAACCAGGTT 60

Qy       1390 GAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGTGAGGATGAT 1449
              |||||||||||||| ||| ||||||||| |||||||| |||||||||| ||||||||||
Db         61 GAGGATTCTAGAAGCGACAAAGTTCATCATGAAAATGTCGCTGATCCCTATGAGGATGAT 120

Qy       1450 AGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGCAACCATACA 1509
              |||||||||||||||     |||||||||||||||||||||||||||| ||||| |||
Db        121 AGAAGTACCATCCCGG------TGGAAGTAAGCATGGATATTCCTGTTAGAAACCAGACA 174

Qy       1510 GTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATACTCTGATGTT 1569
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATGCTCTGATGTT 234

Qy       1570 GTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAG---AACTGGA 1626
              |||||||||||||||||||||||||||||||||||||||||||||||||   |||||||
Db        235 GTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGAACAACTGGA 294

Qy       1627 AGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTGACACCC 1686
              | |||||||||||||||  |||||||||||||||||||||| | ||||||||||||||
Db        295 AATGTGCATCACACAGTTGTGCATCCAGCTGGGAATTTGAGCAATATAAAAGTGACACCC 354

Qy       1687 ACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACAAATATG 1746
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACAAACATG 414

Qy       1747 CATAAGACAGATGTCT-GTCAGCATGTATCAGAAATCTCCACACAG 1791
              |||||||||||||||| |||||||||||||||||||||| ||||||
Db        415 CATAAGACAGATGTCTGGTCAGCATGTATCAGAAATCTCAACACAG 460


RESULT 10
CW686090/c
LOCUS       CW686090                384 bp    DNA     linear   GSS 01-NOV-2004
DEFINITION  OG_BBa0044H01.r OG_BBa Oryza glaberrima genomic clone OG_BBa0044H01
            3', genomic survey sequence.
ACCESSION   CW686090
VERSION     CW686090.1  GI:55155926
KEYWORDS    GSS.
SOURCE      Oryza glaberrima (African rice)
  ORGANISM  Oryza glaberrima
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Kim,H., Yu,Y., Wissotski,M., Byrne,M., Stum,D., Smart,D., Rao,K.,
            Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C.
            and Wing,R.
  TITLE     OMAP
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
```

```
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0044  row: H  column: 01
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .384
                      /organism="Oryza glaberrima"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4538"
                      /clone="OG_BBa0044H01"
                      /tissue_type="young leaves"
                      /lab_host="DH10B T1 phage resistant"
                      /clone_lib="OG_BBa"
                      /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             9.3%;  Score 361.8;  DB 10;  Length 384;
  Best Local Similarity   97.9%;  Pred. No. 1.2e-89;
  Matches  377;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       2437 CATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCA 2496
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| ||
Db        384 CATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATT-GGGCCGTAAGGACACA 326

Qy       2497 AAGAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGT 2556
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AAGAAGCTAACGTGAGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGT 266

Qy       2557 GGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGA 2616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGA 206

Qy       2617 TCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAA 2676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAA 146

Qy       2677 AGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCG 2736
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        145 AGAGCGGTTAACCAGGTCCATGCAAGAAAGTTTCCAAGCACAATAGCAACCATGGAAGCG 86

Qy       2737 AGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATG 2796
             |||||||||||||||||||||||||||||||||||||||||||||||||| || ||||| |
Db          85 AGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTACAGCATCCACG 26

Qy       2797 CCTGCGACGCATCTTCTGAGAATGA 2821
              || ||||||||||||||||||||||
Db          25 CCAGCGACGCATCTTCTGAGAATGA 1


RESULT 11
BZ795917/c
LOCUS       BZ795917               1025 bp    DNA     linear   GSS 17-MAR-2003
DEFINITION  PUFHM45TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa320H17,
            genomic survey sequence.
ACCESSION   BZ795917
VERSION     BZ795917.1  GI:28996773
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUFHM45TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES              Location/Qualifiers
     source           1. .1025
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /strain="B73"
                      /db_xref="taxon:4577"
                      /clone="ZMMBTa320H17"
                      /clone_lib="ZM_0.6_1.0_KB"
                      /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                      CoT selected genomic DNA library"
ORIGIN

  Query Match             9.1%;  Score 355.6;  DB 9;  Length 1025;
  Best Local Similarity   67.0%;  Pred. No. 1e-87;
  Matches  619;  Conservative    0;  Mismatches  254;  Indels   51;  Gaps    6;

Qy       2339 CTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAG 2398
             ||   | |  | | |  ||  ||| ||| || ||||| |||||||| || |  | ||| |
Db        888 CTGTCCTTGATGACATTTTGGAGTGTACCCAAGAACACCAGACACAATTCCACAGGGATG 829

Qy       2399 AAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTG 2458
             |  |||||| |||||     | ||||||  | ||||||||||      ||||  ||||||
Db        828 GAGGGGTCACCATTGCATGTACCTCACCTATGTTTTCACATC------ATCAACATATTG 775

Qy       2459 CTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAAT 2518
             ||||||  ||  ||      ||| |   ||| |  | ||||||| ||| ||||||    |
Db        774 CTGAAGTGCCTGCTCGGAGCTGGAGGAATAACGGGGAAAAGAAGTTAATGTGGGATTCTT 715

Qy       2519 TTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTA 2578
             | ||| |  || ||||||||| ||| || |  ||| ||||  ||||||  ||     |
Db        714 TCAAGACAGCTTCAAGAAATTCACCAACATCGTCATATGGTTTTCAATTCGGAAACAGGC 655

Qy       2579 TCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTC 2638
             | |||| |||| | |   || | | ||||      ||| |||| || || |||||| |||
Db        654 TTCAAGAAGTTCATTCATCTCCCAGTCATGCTTATGGAGCTTCTAGTAACTATGCAGCTC 595

Qy       2639 GCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATG 2698
             |||  |||||||||| | |  | ||                                ||
Db        594 ACCAGCCAGTAATTGTGGCTGTAGA----------------------------TCAGC 565

Qy       2699 CAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAA 2758
             | |||| | ||||||||||| | |  || ||||| ||  ||| ||| | |||||| || |
Db        564 CGAGAAGTGTTCCAAGCACAACATCGACTATGGAGGCTGGTAGGTTGTATGATCGAAGGA 505

Qy       2759 ATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAA 2818
             | |||||||   || | | || || |||||  | ||||||| || ||||||||||||
Db        504 TTCCTGGACAGTCAGGCCTCTACCCAAAAGAGCCTATGCCTGCAACACATCTTCTGAGAT 445

Qy       2819 TGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGA 2878
             |||||||  ||||||||     ||  |||| | ||||||  ||   ||| ||||  |||||
Db        444 TGATGGACTCATCAACAGCTCCAGACTTCACAAACTATCAAAGAGCTAACAGGCGCCAGA 385

Qy       2879 TGGAGTCTCAACTTCATA------ATTCTCAGTATGCAC---ATAATCAGTACAAAGGAT 2929
             ||||   ||||   || |      |||| || ||| |||   || ||||||| || || |
Db        384 TGGAACATCAAACACAAACTCTGGATTCACAATATACACAGCATGATCAGTATAATGGGT 325

Qy       2930 CAACCAGCACATCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACT 2989
             ||   |||||||||||||| || || | | ||||||| |||||  |||| |  ||||||
Db        324 CACTGAGCACATCATATGGAAGACACATAACTGGAAAGGTTCCACCGACACTGCAAGACT 265

Qy       2990 TATCACGGCATCAG---CTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTG 3046
             |||| |||| ||||    | |||  | |||||||| |||||||||| |||||| | || |
Db        264 TATCTCGGCGTCAGGTCCAGCAAAACCTGCACAGGCCTTTACGCCCTCATCCTCGGGTGG 205

Qy       3047 GTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGG---CA 3103
             | |||||||| || ||||||||| |||| || ||||||||||  || |||||  ||
```

```
Db        204 GCGTGCTTGGTTCGTTGCTGCAGCAGGATATCGCAAACTGGTATGAAAACTGCGGGCCGC 145

Qy       3104 CACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAA 3163

Db        144 CACAGTCTGGATACAGGCTGGGCGTTTCTAAAGGGACGGCATCGCTTGATATGAACAGAC 85

Qy       3164 AGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGT 3223

Db         84 ATGGAAACTACGAGACCTTGAACTCAGGAGTGTTCACAGCAGGATGGAATGCCCTTCAAT 25

Qy       3224 TGGGTTCTGTTAGCTCCAGTGCAG 3247

Db         24 TGGGTTCTGTTAGCTCTCTTGCCG 1


RESULT 12
CC355521
LOCUS       CC355521                888 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  PUHPL93TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa504O18,
            genomic survey sequence.
ACCESSION   CC355521
VERSION     CC355521.1  GI:30824921
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHPL93TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa504O18"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             9.0%;  Score 352;  DB 9;  Length 888;
  Best Local Similarity   67.6%;  Pred. No. 9.7e-87;
  Matches  607;  Conservative    0;  Mismatches  240;  Indels   51;  Gaps    6;

Qy       2538 TTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGAC 2597

Db          1 TTCACCAACATCGTCATATGGTTTTCAATTCGGAAACAGGCTTCAAGAAGTTCATTCATC 60

Qy       2598 TTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCC 2657

Db         61 TCCCAGTCATGCTTATGGAGCTTCTAGTAACTATGCAGCTCACCAGCCAGTAATTGTGGC 120

Qy       2658 ACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCAC 2717

Db        121 TATAGAT----------------------------CAGCCGAGAAGTGTTCCAAGCAC 150

Qy       2718 AATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTT 2777

Db        151 AACATCGACTATGGAGGCTGGTAGGTTGTATGATCGAAGGATTCCTGGACAGTCAGGCCT 210

Qy       2778 GTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATT 2837

Db        211 CTACCCAAAAGAGCCTATGCCTGCAACACATCTTCTGAGATTGATGGACTCATCAACAGC 270

Qy       2838 AGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATA- 2896

Db        271 TCCAGACTTCACAAACTATCAAAGAGCTAACAGGCGCCAGATGGAACATCAAACACAAAC 330

Qy       2897 -----ATTCTCAGTATGCA---CATAATCAGTACAAAGGATCAACCAGCACATCATATGG 2948

Db        331 TCTGGATTCACAATATACACAGCATGATCAGTATAATGGGTCACTGAGCACATCATATGG 390

Qy       2949 CAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCT--- 3005

Db        391 AAGACACATAACTGGAAAGGTTCCACCGACACTGCAAGACTTATCTCGGCGTCAGGTCCA 450

Qy       3006 GCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCT 3065

Db        451 GCAAAACCTGCACAGGCCTTTACGCCCTCATCCTCGGGTGGGCGTGCTTGGTTCGTTGCT 510

Qy       3066 GCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGG---CACACAATCTGGTTATAAGTT 3122

Db        511 GCAGCAGGATATCGCAAACTGGTATGAAAACTGCGGGCCGCCACAGTCTGGATACAGGCT 570

Qy       3123 AGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCT 3182

Db        571 GGGCGTTTCTAAAGGGACGGCATCGCTTGATATGAACAGACATGGAAACTACGAGACCTT 630

Qy       3183 GAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCT---- 3238

Db        631 GAACTCAGGAGTGTTCACAGCAGGATGGAATGCCCTTCAATTGGGTTCTGTTAGCTCTCT 690

Qy       3239 --CCAGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAA 3296

Db        691 TGCCGATCCAGAGTACCGGTTGCCAAGGTATGGTACAGGTCAGCCTTGTACAGGTGGCAA 750

Qy       3297 GGGTAAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAA 3356

Db        751 TGGGAGAACGGTTCACCCGTTGGATAAGCTTGTGCGAAAGGATATCTGTGTGACTAACAG 810

Qy       3357 GAACCCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGCTGA 3414

Db        811 AAACCCAGCTGATTTTACTGTAATTAGTGACATGAATGAGTATATGAGAAGCCTTTGA 868


RESULT 13
CL704355/c
LOCUS       CL704355                731 bp    DNA     linear   GSS 22-NOV-2004
DEFINITION  SP__Bb0014N03.r SP__Bb Sorghum propinquum genomic clone
            SP__Bb0014N03 3', genomic survey sequence.
ACCESSION   CL704355
VERSION     CL704355.2  GI:55940476
KEYWORDS    GSS.
SOURCE      Sorghum propinquum
  ORGANISM  Sorghum propinquum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Kudrna,D.,
            Muller,C., Soderlund,C., Bowers,J.E., Paterson,A.H. and Wing,R.
  TITLE     Physical mapping of the sorghum genome
  JOURNAL   Unpublished (2004)
COMMENT     On Nov 22, 2004 this sequence version replaced gi:50270630.
            Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
```

```
          Plate: 0014  row: N  column: 03
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Sorghum propinquum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:132711"
                     /clone="SP__Bb0014N03"
                     /lab_host="DH10B"
                     /clone_lib="SP__Bb"
                     /note="Vector: pCUGI1; Site_1: HindIII; Site_2: HindIII;
                     NSF Sorghum propinquum library"
ORIGIN

  Query Match             8.4%;  Score 327;  DB 10;  Length 731;
  Best Local Similarity   68.8%;  Pred. No. 9.6e-80;
  Matches  501;  Conservative    0;  Mismatches  206;  Indels   21;  Gaps    3;

Qy       2614 GGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCT 2673
Db        728 GGAGCCTCTAGTAACTATGCAGCTCACCAGCCAGTAATTGTGGCTGTAGATCCCTACACA 669

Qy       2674 GAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAA 2733
Db        668 AAGGAAGCAGTTAGCCTGGTTCATNCAAGAAATGTCCCAAGCACAGCATTGACTATGGGG 609

Qy       2734 GCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCC 2793
Db        608 TCTGGTAGGCTGTACGATCAAAGGATTCCTGGACAGTCAGGCCTTTATCCAAAAGAGCCT 549

Qy       2794 ATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAAC 2853
Db        548 GTGCCCGCTACACATCTTCTGAGACTGATGGACTCATCAACACCTTCAGGCTTCACAAGC 489

Qy       2854 TATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCA--- 2910
Db        488 TATCAAAGACCCGACAGGCGCCAGATGGAACTTCATCAAACACAAACTCTCGGTTCACAA 429

Qy       2911 ---------CATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAAT 2961
Db        428 TTTACGCAGCATGATCAGTATAATGCATCACCGAGCACATCATATGGAAGCCACATAATT 369

Qy       2962 GGAAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCT---GCATGATCTGCAC 3018
Db        368 GAAGAGGTTCCACTGACACTGCAGGACTTATCTCGGCGTCAGCTCCAGCCAAACTTGCTC 309

Qy       3019 AGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATT 3078
Db        308 AGGCCTGTACGCCCTCATCCTCGAGTTGGCGTGCTTGGTTCATTGCTGCAGCAGGATATT 249

Qy       3079 GCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGA 3138
Db        248 GCAAACTGGTCTGAAAACTATGGGCCACAGTCTGGGTACAGATTAGGTGATTCTAAAGGG 189

Qy       3139 ATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTT 3198
Db        188 ACAACAACGCTTGATATGAACAGAAACGGAAACTACAAGGCCTTGAACTCAGGAATGCTC 129

Qy       3199 TCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCT------CCAGTGCAGATTTT 3252
Db        128 TCAGCAGGATGGAGTGCTCTGCAGTTGGGTTCTGTTACCTCTGTTGCCAATCCAGAGTAC 69

Qy       3253 TTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCAT 3312
Db          68 CAATTGCCAAGGTATGGTACAGGTCAACCTTCTACCAGTGGCAATGGGAAAACAATTCAT 9

Qy       3313 CCCTTGGA 3320
Db          8 CCGTTGGA 1


RESULT 14
CL597614
```

```
LOCUS       CL597614               647 bp    DNA     linear   GSS 10-JUN-2004
DEFINITION  OB__Ba0059J21.r OB__Ba Oryza brachyantha genomic clone
            OB__Ba0059J21 3', genomic survey sequence.
ACCESSION   CL597614
VERSION     CL597614.1  GI:48570923
KEYWORDS    GSS.
SOURCE      Oryza brachyantha
  ORGANISM  Oryza brachyantha
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE     OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 145000   Std Error: 0.00
            Plate: 0059  row: J  column: 21
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Oryza brachyantha"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4533"
                     /clone="OB__Ba0059J21"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OB__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             8.1%;  Score 315.6;  DB 10;  Length 647;
  Best Local Similarity   74.7%;  Pred. No. 1.5e-76;
  Matches  466;  Conservative    0;  Mismatches   69;  Indels   89;  Gaps    2;

Qy        593 CAGATAATGGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTT 652
Db          1 CAGATTATGGAACATCACCAAGAACTCTTCCTGCAAAGCAAGATGGCACAAGCGATGGTT 60

Qy        653 GCTCCATCACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGT 712
Db         61 GCTCCATCACATTTGTTCGGAGCACTTTTATGCCTGCTACTGTTGGTTCGAAAATAGTGT 120

Qy        713 CTCCTAGCACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTG 772
Db        121 CTCCTAGTGCACAATCATCTCAAGGGAAGAAGTCTGATAGATCAACTCTTCCGAAGAGTG 180

Qy        773 TGCAAGAAGGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGG 832
Db        181 CGCAAGAAGGCAACAACTCCAAATGCAATGCACCTTCTGGCCAGAAAGGAGCTGCTGAGG 240

Qy        833 CCAATACTGATTCACCAATGAA---------------------------------------- 854
Db        241 GCAATACTGGTCCACACATGAAAGGTATGGTCGAGGTAGAGCTTTTTAAATTCCTAGATA 300

Qy        855 ----------------------------------------------AGATTTGCAAGG 866
Db        301 GGTTTTTATTTAAGGTATAGAATAAACTGATGTTTGTGTGATTCCCACAGATTTGCAAGG 360

Qy        867 GCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGT 926
```

```
Db        361 GGCAGCCCAAAATTATGATATGGCAGTAAATATCCCT---GACAACACTTCTATAGATGT 417

Qy        927 TGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCA 986

Db        418 TGGAGTGTTGCCTGAAGTTCCCCAGATTGCATTGCACAAAGAAGGGAACGATGAAGATCA 477

Qy        987 ACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGG 1046

Db        478 ATCCCCATCCACTCCAAAACTTTCTGAAGTGATCCTCAAACCAAATGATGATAGAGATGG 537

Qy       1047 AAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAAT 1106

Db        538 AAAAACTAAAGGACTTCTTGTTGCTGAGAAGTGCAATTTGACCAAACATCTGAAACCAAT 597

Qy       1107 GTCTGGAAAGGAACGTGATCAGGT 1130

Db        598 GTCTGGACAGAAATGCGATCAGGT 621


RESULT 15
AG209720
LOCUS       AG209720                491 bp    DNA     linear   GSS 09-AUG-2003
DEFINITION  Oryza sativa (japonica cultivar-group) DNA, clone:NE8014_0_704_1A,
            3' flanking sequence of Tos17 insertion in rice strain NE8014,
            genomic survey sequence.
ACCESSION   AG209720
VERSION     AG209720.1  GI:32356910
KEYWORDS    GSS.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Miyao,A., Tanaka,K., Murata,K., Sawaki,H., Takeda,S., Abe,K.,
            Shinozuka,Y., Onosato,K. and Hirochika,H.
  TITLE     Target Site Specificity of the Tos17 Retrotransposon Shows a
            Preference for Insertion within Genes and against Insertion in
            Retrotransposon-Rich Regions of the Genome
  JOURNAL   Plant Cell 15 (8), 1771-1780 (2003)
   PUBMED   12897251
REFERENCE   2  (bases 1 to 491)
  AUTHORS   Miyao,A., Onosato,K. and Hirochika,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Akio Miyao, National Institute of
            Agrobiological Sciences, Molecular Genetics; 2-1-2, Kannondai,
            Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyao@affrc.go.jp,
            URL:http://tos.nias.affrc.go.jp/, Tel:81-298-38-7020,
            Fax:81-298-38-7020)
FEATURES             Location/Qualifiers
     source          1. .491
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /strain="NE8014"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="NE8014_0_704_1A"
                     /clone_lib="PCR product directly amplified from rice
                     genomic DNA"
                     /note="The 3' end of retrotransposon Tos17 was found
                     immediately upstream of this sequence."
     misc_feature    1
                     /note="This insertion point is base 139 in AF326768."
ORIGIN

  Query Match             6.9%;  Score 268.6;  DB 10;  Length 491;
  Best Local Similarity   93.5%;  Pred. No. 2.2e-63;
  Matches  288;  Conservative    0;  Mismatches   19;  Indels    1;  Gaps    1;

Qy        139 GTGCTAATTCGCGGCGTTACAAGATCACTGCTGGATTGATATTGAGTTGTGCCTCGGCTG 198

Db          1 GTGCTAATTCNCGGCNTTACANGATCACTGCTGGATTGATATTGAGTTGTGCCTCGGCTG 60

Qy        199 TGCTGGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGATATGGAGATTGTTGCAGTA 258

Db         61 TGCTAGCTGTGTGTTGATTCTCTCCTCGTCGTGGTGATCGATATGGAGATTGTTGCAGTA 120

Qy        259 -GATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTTGCTCGTGGTGGAACTGG 317

Db        121 TGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTTGCTCGTGGTGGAACTGG 180

Qy        318 TGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAGGATGCAGCCGCTGAAGC 377

Db        181 TGCTGTACCGCCAGNGTTGGAGCTGACAGCGACACCTCGTCAGGATGCAGCCGCTGAAGC 240

Qy        378 TGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGAGGGTATGTTGC 437

Db        241 TGGTGTANACGAACCGGCACAACACCAATGCGAGCATTTCTCCATAAGGTAATCATTTTC 300

Qy        438 TCTTCTTC 445

Db        301 TGTATTTC 308


Search completed: March  7, 2006, 22:02:02
Job time : 14351 secs
```

GenCore version 5.1.7

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2006, 13:21:16 ; Search time 11523 Seconds
(without alignments)
5214.226 Million cell updates/sec

Title: US-09-828-068-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTNCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US09828068/runat_06032006_163153_25124/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06p
-USER=US09828068_@CGN_1_1_7415_@runat_06032006_163153_25124 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5526	100.0	4310	6	BD083699	BD083699 Novel gen
2	5516	99.8	3896	15	AF326768	AF326768 Oryza sat
3	5191	93.9	9455	6	BD083700	BD083700 Novel gen
4	5191	93.9	110000	15	AP008207_071	Continuation (72 o
5	5191	93.9	150594	15	AP001859	AP001859 Oryza sat
6	5179	93.7	3626	15	AK103364	AK103364 Oryza sat
7	393	7.1	3291	6	AX590041	AX590041 Sequence
8	298	5.4	8648	6	AR368885	AR368885 Sequence
9	298	5.4	8648	15	AF319968	AF319968 Arabidops
10	298	5.4	17341	6	AR368884	AR368884 Sequence
11	298	5.4	101665	15	ATF15N18	AL163815 Arabidops
c 12	269	4.9	108174	15	AP006698	AP006698 Lotus cor
13	251	4.5	5397	11	AY605097	AY605097 Synthetic
14	248	4.5	3144	5	CR848207	CR848207 Xenopus t
15	246	4.5	3154	5	BC064870	BC064870 Xenopus t
16	241.5	4.4	7036	9	MUSCYCLORA	L04289 Mus musculu
17	240	4.3	209201	9	AC139244	AC139244 Mus muscu
18	234.5	4.2	32367	9	AF520421S2	AF520422 Mus muscu
19	234.5	4.2	140982	5	BX842684	BX842684 Zebrafish
20	233	4.2	179931	2	AC009256	AC009256 Drosophil
21	233	4.2	198967	2	AC010919	AC010919 Drosophil
22	233	4.2	323840	2	AE003501	AE003501 Drosophil
23	227.5	4.1	212955	14	AC136676	AC136676 Rattus no
c 24	223.5	4.0	160827	2	AC099025	AC099025 Drosophil
25	223.5	4.0	193262	2	AC007579	AC007579 Drosophil
c 26	223.5	4.0	260367	2	AE003808	AE003808 Drosophil
27	223	4.0	3114	8	HSATRX10	U72909 Homo sapien
28	223	4.0	4327	6	CS047818	CS047818 Sequence
29	223	4.0	4808	2	AY075323	AY075323 Drosophil
30	223	4.0	6831	6	CQ718057	CQ718057 Sequence
31	223	4.0	7479	8	AB102641	AB102641 Homo sapi
32	223	4.0	10216	8	HSU72938	U72938 Homo sapien
33	223	4.0	10330	6	CS032835	CS032835 Sequence
34	223	4.0	10330	6	CS041787	CS041787 Sequence
35	223	4.0	10330	6	AX828379	AX828379 Sequence
36	223	4.0	10330	8	HSU72937	U72937 Homo sapien
37	223	4.0	10452	6	AX336451	AX336451 Sequence
38	223	4.0	10452	6	AX828404	AX828404 Sequence
39	223	4.0	10452	8	HSU72936	U72936 Homo sapien
c 40	223	4.0	47323	8	AC005937	AC005937 Homo sapi
41	223	4.0	84474	8	AL662854	AL662854 Human DNA
42	223	4.0	103780	8	AL773541	AL773541 Human DNA
c 43	223	4.0	110000	14	BX276116_03	Continuation (4 of
44	223	4.0	114883	8	CR759747	CR759747 Human DNA
45	223	4.0	135982	8	HS346O6	Z84487 Human DNA s

ALIGNMENTS

RESULT 1
BD083699
LOCUS BD083699 4310 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel gene participating in response to brassinosteroid.
ACCESSION BD083699
VERSION BD083699.1 GI:22629309
KEYWORDS JP 2001327287-A/1.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 4310)
AUTHORS Hirochika,H., Yamazaki,T. and Miyao,A.
TITLE Novel gene participating in response to brassinosteroid
JOURNAL Patent: JP 2001327287-A 1 27-NOV-2001;
NATL INST OF AGROBIOLOGICAL RESOUCES, BIO ORIENTED TECHNOLOGY
RESEARCH ADVA JAPAN FORAGE SEED ASSOCIATION
COMMENT OS Oryza sativa (rice)
PN JP 2001327287-A/1
PD 27-NOV-2001
PF 19-MAY-2000 JP 2000149106
PI HIROHIKO HIROCHIKA,TOKIO YAMAZAKI,AKIO MIYAO
PC C12N15/09,C12N15/00
CC Novel gene participating in response to brassinosteroid FH
Key Location/Qualifiers
FT CDS (655). .(3828).

```
FEATURES             Location/Qualifiers
     source          1. .4310
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
ORIGIN

Alignment Scores:
Pred. No.:                3.3e-287        Length:          4310
Score:                    5526.00         Matches:         1057
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-09-828-068-2 (1-1057) x BD083699 (1-4310)

Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 714

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 774

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 834

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 894

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 954

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1014

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1074

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1134

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1194

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1254

Qy        201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1314

Qy        221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1374

Qy        241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1434

Qy        261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1494

Qy        281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1554

Qy        301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1614

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1674

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1734

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1735 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1794

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1795 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1854

Qy        401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1914

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1974

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2034

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2094

Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2154

Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2214

Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2274

Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2334

Qy        561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2394

Qy        581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2454

Qy        601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2514

Qy        621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2574

Qy        641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
```

```
Db       2575 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2634

Qy        661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680

Db       2635 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2694

Qy        681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700

Db       2695 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2754

Qy        701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720

Db       2755 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2814

Qy        721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740

Db       2815 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2874

Qy        741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760

Db       2875 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2934

Qy        761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780

Db       2935 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2994

Qy        781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800

Db       2995 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3054

Qy        801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820

Db       3055 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3114

Qy        821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840

Db       3115 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3174

Qy        841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860

Db       3175 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3234

Qy        861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880

Db       3235 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3294

Qy        881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900

Db       3295 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3354

Qy        901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920

Db       3355 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3414

Qy        921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940

Db       3415 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3474

Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960

Db       3475 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3534

Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980

Db       3535 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3594

Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000

Db       3595 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3654

Qy       1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
```

```
Db       3655 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3714

Qy       1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040

Db       3715 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3774

Qy       1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057

Db       3775 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3825
```

RESULT 2
AF326768

LOCUS	AF326768 3896 bp mRNA linear PLN 05-SEP-2001
DEFINITION	Oryza sativa embryonic flower 1-like protein mRNA, complete cds.
ACCESSION	AF326768
VERSION	AF326768.1 GI:15430698
KEYWORDS	.
SOURCE	Oryza sativa
ORGANISM	Oryza sativa Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1 (bases 1 to 3896)
AUTHORS	Aubert,D., Chen,L., Moon,Y.H., Martin,D., Castle,L.A., Yang,C.H. and Sung,Z.R.
TITLE	EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis
JOURNAL	Plant Cell 13 (8), 1865-1875 (2001)
PUBMED	11487698
REFERENCE	2 (bases 1 to 3896)
AUTHORS	Moon,Y.-H., Chen,L. and Sung,Z.R.
TITLE	Direct Submission
JOURNAL	Submitted (06-DEC-2000) Plant and Microbial Biol. Dept., University of California, Berkeley, 361 Koshland Hall, Berkeley, CA 94720, USA

FEATURES Location/Qualifiers

source 1. .3896
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"

CDS 241. .3414
/note="EMF1"
/codon_start=1
/product="embryonic flower 1-like protein"
/protein_id="AAK98529.1"
/db_xref="GI:15430699"
/translation="MEIVAVDQEGARVVGTNCMLARGGTGAVAPVLELTATPRQDAAA
EAGVDEPAQHQCEHFSIRGYVALLQKKDPKFCSLSRIFHDQKKCDEHKASSSPFSVAK
FRRWDCSKCLDKLKTSDNGTAPRTLPAKQNGTSDGCSITFVRSTFVPASVGSQKVSPS
TQSSQGKNADRSTLPKSVQEGNDSKCNAPSGKNGAAEANTDSPMKDLQGPAQNYDVAA
NVSEDNTSVDVGALPEVPQITWHIEVNGADQPPSTPKLSEVVLKRNEDENGKTEETLV
AEQCNLTKDPNPMSGKERDQVAEQCNLTKDPKPVSGQKCEQICNEPCEEVVLKRSSKS
KRKTDKKLMKKQQHSKKRTAQADVSDAKLCRRKPKKVRLLSEIVNANQVEDSRSDEVH
RENAADPCEDDRSTIPVPMEVSMDIPVSNHTVGEDGLKSSKNKTKRKYSDVVDDGSSL
MNWLNGKKKRTGSVHHTVAHPAGNLSNKKVTPTASTQHDDENDTENGLDTNMHKTDVC
QHVSEISTQRCSSKGKTAGLSKGKTHSAASTKYGGESTRNGQNIHVLSAEDQCQMETE
NSVLSHSAKVSPAEHDIQIMSDLHEQSLPKKKKKQKLEVTREKQTMIDDIPMDIVELL
AKNQHERQLMTETDCSDINRIQSKTTADDDCVIVAAKDGSDYASSVFDTNSQQKSLAS
QSTQKELQGHLALTTQESPHPQNFQSTQEQQTHLRMEEMVTIAASSPLFSHHDDQYIA
EAPTEHWGRKDAKKLTWEQFKATTRNSPAATCGAQFRPGIQAVDLTSTHVMGSSSNYA
SRQPVIAPLDRYAERAVNQVHARNFPSTIATMEASKLCDRRNAGQVVLYPKESMPATH
LLRMMDPSTLASFPNYGTSSRNQMESQLHNSQYAHNQYKGSTSTSYGSNLNGKIPLTF
EDLSRHQLHDLHRPLRPHPRVGVLGSLLQKEIANWSENCGTQSGYKLGVSTGITSHQM
NRKEHFEALNSGMFSAKWNALQLGSVSSSAYFLSARNSIAQSWTRGKGKMVHPLDRFV
RQDICITNKNPADFTTISNDNEYMDYR"

ORIGIN

Alignment Scores:

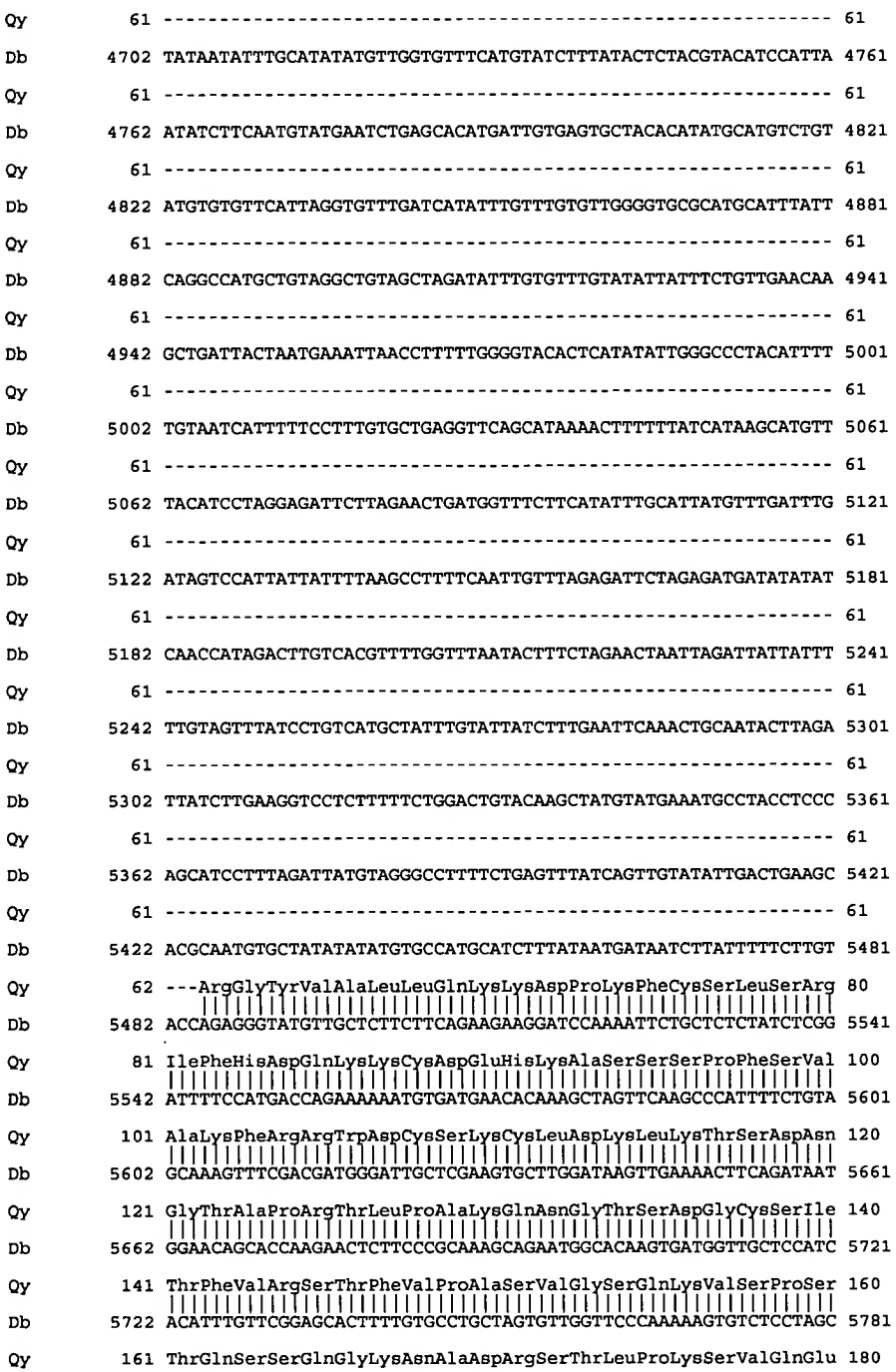
Pred. No.:	1.01e-286	Length:	3896
Score:	5516.00	Matches:	1055
Percent Similarity:	99.9%	Conservative:	1
Best Local Similarity:	99.8%	Mismatches:	1
Query Match:	99.8%	Indels:	0
DB:	15	Gaps:	0

US-09-828-068-2 (1-1057) x AF326768 (1-3896)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840

Qy        201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900

Qy        221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960

Qy        241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020

Qy        261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080

Qy        281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140

Qy        301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTGTAAATGCT 1380

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440

Qy        401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680

Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740

Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800

Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860

Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920

Qy        561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980

Qy        581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040

Qy        601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100

Qy        621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160

Qy        641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220

Qy        661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280

Qy        681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340

Qy        701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
```

```
Qy        721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460

Qy        741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520

Qy        761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580

Qy        781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640

Qy        801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700

Qy        821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760

Qy        841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820

Qy        861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880

Qy        881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940

Qy        901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000

Qy        921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060

Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120

Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180

Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240

Qy       1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGTGCATATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300

Qy       1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360

Qy       1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3411

RESULT 3
BD083700
LOCUS       BD083700                9455 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Novel gene participating in response to brassinosteroid.
ACCESSION   BD083700
VERSION     BD083700.1  GI:22629310
KEYWORDS    JP 2001327287-A/2.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 9455)
  AUTHORS   Hirochika,H., Yamazaki,T. and Miyao,A.
  TITLE     Novel gene participating in response to brassinosteroid
  JOURNAL   Patent: JP 2001327287-A 2 27-NOV-2001;
            NATL INST OF AGROBIOLOGICAL RESOUCES, BIO ORIENTED TECHNOLOGY
            RESEARCH ADVANCEMENT INSTITUTION
COMMENT     OS   Oryza sativa (rice)
            PN   JP 2001327287-A/2
            PD   27-NOV-2001
            PF   19-MAY-2000 JP 2000149106
            PI   HIROHIKO HIROCHIKA,TOKIO YAMAZAKI,AKIO MIYAO
            PC   C12N15/09,C12N15/00
            CC   Novel gene participating in response to brassinosteroid FH
            Key             Location/Qualifiers
            FT   source          1. .9455
            FT                   /organism='Oryza sativa (rice)'.
FEATURES             Location/Qualifiers
     source          1. .9455
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
ORIGIN

Alignment Scores:
Pred. No.:              7.88e-269       Length:         9455
Score:                  5191.00         Matches:        1055
Percent Similarity:     68.1%           Conservative:   0
Best Local Similarity:  68.1%           Mismatches:     2
Query Match:            93.9%           Indels:         495
DB:                     6               Gaps:           3

US-09-828-068-2 (1-1057) x BD083700 (1-9455)

Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4223 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 4282

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4283 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 4342

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4343 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 4402

Qy         61 Ile------------------------------------------------------------ 61

Db       4403 AT-AAGGTAATCATTTTCTGTATTTCCAATTCCAGTATCGCGTTGTGGATGAATAATGAA 4461

Qy         61 ------------------------------------------------------------ 61

Db       4462 TCGGCATGTCATGCCATATTGCACTGTTTGATGGAAGAGTATGATTGATCGTGGTTTTTG 4521

Qy         61 ------------------------------------------------------------ 61

Db       4522 CACAGTTTGCTGTTGGGACTTATATGGTCATCTGTTTTGTACGATCGTATACACTGGGTC 4581

Qy         61 ------------------------------------------------------------ 61

Db       4582 GACATGCTTATGACTTTGGTTCGATTTAGGAAGTCAATACATCCACTACTAGCTCTATAT 4641

Qy         61 ------------------------------------------------------------ 61

Db       4642 CTAGCCATGTGAACTCATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAAAAAATATA 4701
```


Qy 61 ---------- 61
Db 4702 TATAATATTTGCATATATGTTGGTGTTTCATGTATCTTTATACTCTACGTACATCCATTA 4761
Qy 61 ---------- 61
Db 4762 ATATCTTCAATGTATGAATCTGAGCACATGATTGTGAGTGCTACACATATGCATGTCTGT 4821
Qy 61 ---------- 61
Db 4822 ATGTGTGTTCATTAGGTGTTTGATCATATTTGTTTGTGTTGGGGTGCGCATGCATTTATT 4881
Qy 61 ---------- 61
Db 4882 CAGGCCATGCTGTAGGCTGTAGCTAGATATTTGTGTTTGTATATTATTTCTGTTGAACAA 4941
Qy 61 ---------- 61
Db 4942 GCTGATTACTAATGAAATTAACCTTTTTGGGGTACACTCATATATTGGGCCCTACATTTT 5001
Qy 61 ---------- 61
Db 5002 TGTAATCATTTTTCCTTTGTGCTGAGGTTCAGCATAAAACTTTTTTATCATAAGCATGTT 5061
Qy 61 ---------- 61
Db 5062 TACATCCTAGGAGATTCTTAGAACTGATGGTTTCTTCATATTTGCATTATGTTTGATTTG 5121
Qy 61 ---------- 61
Db 5122 ATAGTCCATTATTATTTTAAGCCTTTTCAATTGTTTAGAGATTCTAGAGATGATATATAT 5181
Qy 61 ---------- 61
Db 5182 CAACCATAGACTTGTCACGTTTTGGTTTAATACTTTCTAGAACTAATTAGATTATTATTT 5241
Qy 61 ---------- 61
Db 5242 TTGTAGTTTATCCTGTCATGCTATTTGTATTATCTTTGAATTCAAACTGCAATACTTAGA 5301
Qy 61 ---------- 61
Db 5302 TTATCTTGAAGGTCCTCTTTTTCTGGACTGTACAAGCTATGTATGAAATGCCTACCTCCC 5361
Qy 61 ---------- 61
Db 5362 AGCATCCTTTAGATTATGTAGGGCCTTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC 5421
Qy 61 ---------- 61
Db 5422 ACGCAATGTGCTATATATATGTGCCATGCATCTTTATAATGATAATCTTATTTTTCTTGT 5481
Qy 62 ---ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
Db 5482 ACCAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 5541
Qy 81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
Db 5542 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 5601
Qy 101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
Db 5602 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 5661
Qy 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
Db 5662 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 5721
Qy 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
Db 5722 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 5781
Qy 161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
Db 5782 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 5841
Qy 181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
Db 5842 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 5901
Qy 201 AspSerProMetLys---------- 205
Db 5902 GATTCACCAATGAA-AGGTATGGTAGATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTT 5960
Qy 206 ----------AspLeuGlnGlyProAla 211
Db 5961 ATTTAAGGTATAGAATAAACTAATGTTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC 6020
Qy 212 GlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAla 231
Db 6021 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 6080
Qy 232 LeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
Db 6081 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 6140
Qy 252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
Db 6141 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 6200
Qy 272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
Db 6201 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGA 6260
Qy 292 LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
Db 6261 AAGGAACGTGATCAGGTTGCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCT 6320
Qy 312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
Db 6321 GGGCAGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGC 6380
Qy 332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
Db 6381 TCCAAATCTAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAA 6440
Qy 352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
Db 6441 CGCACTGCCCAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTG 6500
Qy 372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
Db 6501 CGGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTT 6560
Qy 392 HisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMet 411
Db 6561 CATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATG 6620
Qy 412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
Db 6621 GAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCA 6680
Qy 432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
Db 6681 AGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATG 6740
Qy 452 AsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisPro 471
Db 6741 AACTGGCTGAATGGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCA 6800
Qy 472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
Db 6801 GCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAG 6860
Qy 492 AsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVal 511

```
Db       6861 AATGATACTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTGTCAGCATGTA 6920
Qy        512 SerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGly 531
Db       6921 TCAGAAATCTCCACACAGAGGTGCTCATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGG 6980
Qy        532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
Db       6981 AAAACACATTCAGCTGCTAGTACCAAATATGGTGGTGAAAGCACCAGAAATGGTCAGAAC 7040
Qy        552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
Db       7041 ATACATGTACTCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGT 7100
Qy        572 HisSerAlaLys---------------------------------------------- 575
Db       7101 CACTCGGCAAAGGTACGAATTTTGTGAATCATGAGGAATTTTTGCTTTTTAAATTGACTG 7160
Qy        575 ---------------------------------------------------------- 575
Db       7161 AATCAACATTTATCTGTATGAAGGAATAATATTGGTGCATAACAATGTTAAGAAATATGC 7220
Qy        575 ---------------------------------------------------------- 575
Db       7221 ATACAATGTTTATTTATATGCTTTCCACTGTTCTTCTTTACTTATGTTTTGATACTCTTT 7280
Qy        575 ---------------------------------------------------------- 575
Db       7281 TTGTGTGTGCGTGCATGTGTGCATGTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGC 7340
Qy        575 ---------------------------------------------------------- 575
Db       7341 ACGTGCGTGGCGCAATATTCTTTTTTAGACTCATATTATAGTGATTGTAATGGACTGACA 7400
Qy        576 ----------------------ValSerProAlaGluHisAspIleGlnIleMetSerAs 588
Db       7401 TTTTCCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGA 7460
Qy        588 pLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGl 608
Db       7461 CCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGA 7520
Qy        608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGl 628
Db       7521 AAAACAGACCATGATAGATGACATCCCCATGGATATTGTTGAACTGCTAGCTAAAAACCA 7580
Qy        628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLy 648
Db       7581 GCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCTGACATCAACCGTATTCAATCCAA 7640
Qy        648 sThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe 668
Db       7641 GACAACTGCTGATGATGATTGTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCATC 7700
Qy        668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGl 688
Db       7701 AAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGA 7760
Qy        688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSe 708
Db       7761 GTTACAGGGTCATTTGGCATTGACCACACAAGAGTCTCCACATCCTCAGAACTTTCAGTC 7820
Qy        708 rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
Db       7821 TACTCAGGAACAGCAGACACATTTGCGGATGGAAGAAATGGTCACTATTGCTGCAAGCTC 7880
Qy        728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGl 748
Db       7881 ACCACTATTTTCACATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGG 7940
Qy        748 yArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
Db       7941 CCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCC 8000

Qy        768 oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh 788
Db       8001 AGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTAC 8060
Qy        788 rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs 808
Db       8061 TCATGTCATGGGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGA 8120
Qy        808 pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnProSerThrIleAl 828
Db       8121 CCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGC 8180
Qy        828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
Db       8181 AACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCC 8240
Qy        848 oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe 868
Db       8241 TAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATTAGCAAG 8300
Qy        868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888
Db       8301 CTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCA 8360
Qy        888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl 908
Db       8361 GTATGCACATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAATGG 8420
Qy        908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
Db       8421 AAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACAGACC 8480
Qy        928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948
Db       8481 TTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAA 8540
Qy        948 nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleTh 968
Db       8541 CTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAAC 8600
Qy        968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
Db       8601 ATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGC 8660
Qy        988 aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr 1008
Db       8661 AAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAG 8720
Qy       1008 gAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
Db       8721 GAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATCG 8780
Qy       1028 gPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSe 1048
Db       8781 GTTTGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATTTTACTACAATCAG 8840
Qy       1048 rAsnAspAsnGluTyrMetAspTyrArg 1057
Db       8841 TAACGATAACGAGTATATGGATTACCGC 8868
```

RESULT 4
AP008207_071
WPCOMMENT
Sequence split into 433 fragments LOCUS AP008207 Accession AP008207

Fragment Name	Begin	End
AP008207_000	1	110000
AP008207_001	100001	210000
AP008207_002	200001	310000
AP008207_003	300001	410000
AP008207_004	400001	510000
AP008207_005	500001	610000
AP008207_006	600001	710000

AP008207_007	700001	810000
AP008207_008	800001	910000
AP008207_009	900001	1010000
AP008207_010	1000001	1110000
AP008207_011	1100001	1210000
AP008207_012	1200001	1310000
AP008207_013	1300001	1410000
AP008207_014	1400001	1510000
AP008207_015	1500001	1610000
AP008207_016	1600001	1710000
AP008207_017	1700001	1810000
AP008207_018	1800001	1910000
AP008207_019	1900001	2010000
AP008207_020	2000001	2110000
AP008207_021	2100001	2210000
AP008207_022	2200001	2310000
AP008207_023	2300001	2410000
AP008207_024	2400001	2510000
AP008207_025	2500001	2610000
AP008207_026	2600001	2710000
AP008207_027	2700001	2810000
AP008207_028	2800001	2910000
AP008207_029	2900001	3010000
AP008207_030	3000001	3110000
AP008207_031	3100001	3210000
AP008207_032	3200001	3310000
AP008207_033	3300001	3410000
AP008207_034	3400001	3510000
AP008207_035	3500001	3610000
AP008207_036	3600001	3710000
AP008207_037	3700001	3810000
AP008207_038	3800001	3910000
AP008207_039	3900001	4010000
AP008207_040	4000001	4110000
AP008207_041	4100001	4210000
AP008207_042	4200001	4310000
AP008207_043	4300001	4410000
AP008207_044	4400001	4510000
AP008207_045	4500001	4610000
AP008207_046	4600001	4710000
AP008207_047	4700001	4810000
AP008207_048	4800001	4910000
AP008207_049	4900001	5010000
AP008207_050	5000001	5110000
AP008207_051	5100001	5210000
AP008207_052	5200001	5310000
AP008207_053	5300001	5410000
AP008207_054	5400001	5510000
AP008207_055	5500001	5610000
AP008207_056	5600001	5710000
AP008207_057	5700001	5810000
AP008207_058	5800001	5910000
AP008207_059	5900001	6010000
AP008207_060	6000001	6110000
AP008207_061	6100001	6210000
AP008207_062	6200001	6310000
AP008207_063	6300001	6410000
AP008207_064	6400001	6510000
AP008207_065	6500001	6610000
AP008207_066	6600001	6710000
AP008207_067	6700001	6810000
AP008207_068	6800001	6910000
AP008207_069	6900001	7010000
AP008207_070	7000001	7110000
AP008207_071	7100001	7210000
AP008207_072	7200001	7310000
AP008207_073	7300001	7410000
AP008207_074	7400001	7510000
AP008207_075	7500001	7610000
AP008207_076	7600001	7710000
AP008207_077	7700001	7810000
AP008207_078	7800001	7910000
AP008207_079	7900001	8010000
AP008207_080	8000001	8110000
AP008207_081	8100001	8210000
AP008207_082	8200001	8310000
AP008207_083	8300001	8410000
AP008207_084	8400001	8510000
AP008207_085	8500001	8610000
AP008207_086	8600001	8710000
AP008207_087	8700001	8810000
AP008207_088	8800001	8910000
AP008207_089	8900001	9010000
AP008207_090	9000001	9110000
AP008207_091	9100001	9210000
AP008207_092	9200001	9310000
AP008207_093	9300001	9410000
AP008207_094	9400001	9510000
AP008207_095	9500001	9610000
AP008207_096	9600001	9710000
AP008207_097	9700001	9810000
AP008207_098	9800001	9910000
AP008207_099	9900001	10010000
AP008207_100	10000001	10110000
AP008207_101	10100001	10210000
AP008207_102	10200001	10310000
AP008207_103	10300001	10410000
AP008207_104	10400001	10510000
AP008207_105	10500001	10610000
AP008207_106	10600001	10710000
AP008207_107	10700001	10810000
AP008207_108	10800001	10910000
AP008207_109	10900001	11010000
AP008207_110	11000001	11110000
AP008207_111	11100001	11210000
AP008207_112	11200001	11310000
AP008207_113	11300001	11410000
AP008207_114	11400001	11510000
AP008207_115	11500001	11610000
AP008207_116	11600001	11710000
AP008207_117	11700001	11810000
AP008207_118	11800001	11910000
AP008207_119	11900001	12010000
AP008207_120	12000001	12110000
AP008207_121	12100001	12210000
AP008207_122	12200001	12310000
AP008207_123	12300001	12410000
AP008207_124	12400001	12510000
AP008207_125	12500001	12610000
AP008207_126	12600001	12710000
AP008207_127	12700001	12810000
AP008207_128	12800001	12910000
AP008207_129	12900001	13010000
AP008207_130	13000001	13110000
AP008207_131	13100001	13210000
AP008207_132	13200001	13310000
AP008207_133	13300001	13410000
AP008207_134	13400001	13510000
AP008207_135	13500001	13610000
AP008207_136	13600001	13710000
AP008207_137	13700001	13810000
AP008207_138	13800001	13910000
AP008207_139	13900001	14010000
AP008207_140	14000001	14110000
AP008207_141	14100001	14210000
AP008207_142	14200001	14310000
AP008207_143	14300001	14410000
AP008207_144	14400001	14510000
AP008207_145	14500001	14610000
AP008207_146	14600001	14710000
AP008207_147	14700001	14810000
AP008207_148	14800001	14910000
AP008207_149	14900001	15010000
AP008207_150	15000001	15110000
AP008207_151	15100001	15210000
AP008207_152	15200001	15310000

AP008207_153	15300001	15410000
AP008207_154	15400001	15510000
AP008207_155	15500001	15610000
AP008207_156	15600001	15710000
AP008207_157	15700001	15810000
AP008207_158	15800001	15910000
AP008207_159	15900001	16010000
AP008207_160	16000001	16110000
AP008207_161	16100001	16210000
AP008207_162	16200001	16310000
AP008207_163	16300001	16410000
AP008207_164	16400001	16510000
AP008207_165	16500001	16610000
AP008207_166	16600001	16710000
AP008207_167	16700001	16810000
AP008207_168	16800001	16910000
AP008207_169	16900001	17010000
AP008207_170	17000001	17110000
AP008207_171	17100001	17210000
AP008207_172	17200001	17310000
AP008207_173	17300001	17410000
AP008207_174	17400001	17510000
AP008207_175	17500001	17610000
AP008207_176	17600001	17710000
AP008207_177	17700001	17810000
AP008207_178	17800001	17910000
AP008207_179	17900001	18010000
AP008207_180	18000001	18110000
AP008207_181	18100001	18210000
AP008207_182	18200001	18310000
AP008207_183	18300001	18410000
AP008207_184	18400001	18510000
AP008207_185	18500001	18610000
AP008207_186	18600001	18710000
AP008207_187	18700001	18810000
AP008207_188	18800001	18910000
AP008207_189	18900001	19010000
AP008207_190	19000001	19110000
AP008207_191	19100001	19210000
AP008207_192	19200001	19310000
AP008207_193	19300001	19410000
AP008207_194	19400001	19510000
AP008207_195	19500001	19610000
AP008207_196	19600001	19710000

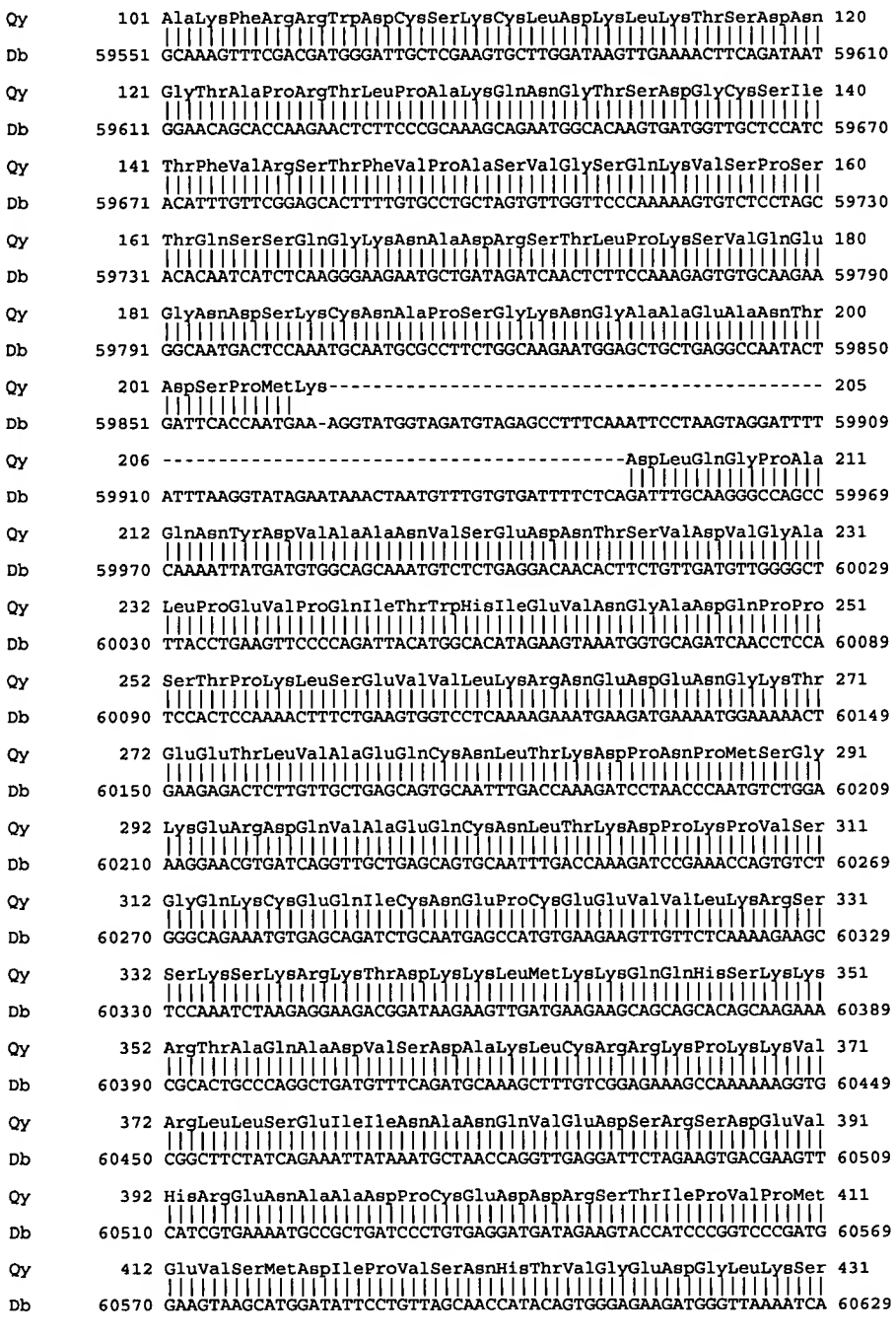
Alignment Scores:

Pred. No.:	1.41e-267	Length:	110000
Score:	5191.00	Matches:	1055
Percent Similarity:	68.1%	Conservative:	0
Best Local Similarity:	68.1%	Mismatches:	2
Query Match:	93.9%	Indels:	495
DB:	15	Gaps:	3

US-09-828-068-2 (1-1057) x AP008207_071 (1-110000)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58172 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 58231

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58232 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 58291

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58292 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 58351

Qy         61 Ile---------------------------------------------------------- 61

Db      58352 AT-AAGGTAATCATTTTCTGTATTTCCAATTCCAGTATCGCGTTGTGGATGAATAATGAA 58410

Qy         61 ---------------------------------------------------------- 61
```

```
Db      58411 TCGGCATGTCATGCCATATTGCACTGTTTGATGGAAGAGTATGATTGATCGTGGTTTTTG 58470

Qy         61 ---------------------------------------------------------- 61

Db      58471 CACAGTTTGCTGTTGGGACTTATATGGTCATCTGTTTTGTACGATCGTATACACTGGGTC 58530

Qy         61 ---------------------------------------------------------- 61

Db      58531 GACATGCTTATGACTTTGGTTCGATTTAGGAAGTCAATACATCCACTACTAGCTCTATAT 58590

Qy         61 ---------------------------------------------------------- 61

Db      58591 CTAGCCATGTGAACTCATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAAAAAATATA 58650

Qy         61 ---------------------------------------------------------- 61

Db      58651 TATAATATTTGCATATATGTTGGTGTTTCATGTATCTTTATACTCTACGTACATCCATTA 58710

Qy         61 ---------------------------------------------------------- 61

Db      58711 ATATCTTCAATGTATGAATCTGAGCACATGATTGTGAGTGCTACACATATGCATGTCTGT 58770

Qy         61 ---------------------------------------------------------- 61

Db      58771 ATGTGTGTTCATTAGGTGTTTGATCATATTTGTTTGTGTTGGGGTGCGCATGCATTTATT 58830

Qy         61 ---------------------------------------------------------- 61

Db      58831 CAGGCCATGCTGTAGGCTGTAGCTAGATATTTGTGTTTGTATATTATTTCTGTTGAACAA 58890

Qy         61 ---------------------------------------------------------- 61

Db      58891 GCTGATTACTAATGAAATTAACCTTTTTGGGGTACACTCATATATTGGGCCCTACATTTT 58950

Qy         61 ---------------------------------------------------------- 61

Db      58951 TGTAATCATTTTTCCTTTGTGCTGAGGTTCAGCATAAAACTTTTTTATCATAAGCATGTT 59010

Qy         61 ---------------------------------------------------------- 61

Db      59011 TACATCCTAGGAGATTCTTAGAACTGATGGTTTCTTCATATTTGCATTATGTTTGATTTG 59070

Qy         61 ---------------------------------------------------------- 61

Db      59071 ATAGTCCATTATTATTTTAAGCCTTTTCAATTGTTTAGAGATTCTAGAGATGATATATAT 59130

Qy         61 ---------------------------------------------------------- 61

Db      59131 CAACCATAGACTTGTCACGTTTTGGTTTAATACTTTCTAGAACTAATTAGATTATTATTT 59190

Qy         61 ---------------------------------------------------------- 61

Db      59191 TTGTAGTTTATCCTGTCATGCTATTTGTATTATCTTTGAATTCAAACTGCAATACTTAGA 59250

Qy         61 ---------------------------------------------------------- 61

Db      59251 TTATCTTGAAGGTCCTCTTTTTCTGGACTGTACAAGCTATGTATGAAATGCCTACCTCCC 59310

Qy         61 ---------------------------------------------------------- 61

Db      59311 AGCATCCTTTAGATTATGTAGGGCCTTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC 59370

Qy         61 ---------------------------------------------------------- 61

Db      59371 ACGCAATGTGCTATATATATGTGCCATGCATCTTTATAATGATAATCTTATTTTTCTTGT 59430

Qy         62 ---ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59431 ACCAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 59490

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59491 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 59550
```


Qy 101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
Db 59551 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 59610
Qy 121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
Db 59611 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 59670
Qy 141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
Db 59671 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 59730
Qy 161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
Db 59731 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 59790
Qy 181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
Db 59791 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 59850
Qy 201 AspSerProMetLys---------- 205
Db 59851 GATTCACCAATGAA-AGGTATGGTAGATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTT 59909
Qy 206 ----------AspLeuGlnGlyProAla 211
Db 59910 ATTTAAGGTATAGAATAAACTAATGTTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC 59969
Qy 212 GlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAla 231
Db 59970 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 60029
Qy 232 LeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
Db 60030 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 60089
Qy 252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
Db 60090 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 60149
Qy 272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
Db 60150 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGA 60209
Qy 292 LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
Db 60210 AAGGAACGTGATCAGGTTGCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCT 60269
Qy 312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
Db 60270 GGGCAGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGC 60329
Qy 332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
Db 60330 TCCAAATCTAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAA 60389
Qy 352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
Db 60390 CGCACTGCCCAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTG 60449
Qy 372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
Db 60450 CGGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTT 60509
Qy 392 HisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMet 411
Db 60510 CATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATG 60569
Qy 412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
Db 60570 GAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCA 60629
Qy 432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
Db 60630 AGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATG 60689
Qy 452 AsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisPro 471
Db 60690 AACTGGCTGAATGGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCA 60749
Qy 472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
Db 60750 GCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAG 60809
Qy 492 AsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVal 511
Db 60810 AATGATACTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTGTCAGCATGTA 60869
Qy 512 SerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGly 531
Db 60870 TCAGAAATCTCCACACAGAGGTGCTCATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGG 60929
Qy 532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
Db 60930 AAAACACATTCAGCTGCTAGTACCAAATATGGTGGTGAAAGCACCAGAAATGGTCAGAAC 60989
Qy 552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
Db 60990 ATACATGTACTCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGT 61049
Qy 572 HisSerAlaLys---------- 575
Db 61050 CACTCGGCAAAGGTACGAATTTTGTGAATCATGAGGAATTTTTGCTTTTTAAATTGACTG 61109
Qy 575 ---------- 575
Db 61110 AATCAACATTTATCTGTATGAAGGAATAATATTGGTGCATAACAATGTTAAGAAATATGC 61169
Qy 575 ---------- 575
Db 61170 ATACAATGTTTATTTATATGCTTTCCACTGTTCTTCTTTACTTATGTTTTGATACTCTTT 61229
Qy 575 ---------- 575
Db 61230 TTGTGTGTGCGTGCATGTGTGCATGTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGC 61289
Qy 575 ---------- 575
Db 61290 ACGTGCGTGGCGCAATATTCTTTTTTAGACTCATATTATAGTGATTGTAATGGACTGACA 61349
Qy 576 ----------ValSerProAlaGluHisAspIleGlnIleMetSerAs 588
Db 61350 TTTTCCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGA 61409
Qy 588 pLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGl 608
Db 61410 CCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGA 61469
Qy 608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGl 628
Db 61470 AAAACAGACCATGATAGATGACATCCCCATGGATATTGTTGAACTGCTAGCTAAAAACCA 61529
Qy 628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLy 648
Db 61530 GCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCTGACATCAACCGTATTCAATCCAA 61589
Qy 648 sThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe 668
Db 61590 GACAACTGCTGATGATGATTGTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCATC 61649
Qy 668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGl 688
Db 61650 AAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGA 61709
Qy 688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSe 708

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61710 GTTACAGGGTCATTTGGCATTGACCACACAAGAGTCTCCACATCCTCAGAACTTTCAGTC 61769

Qy        708 rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61770 TACTCAGGAACAGCAGACACATTTGCGGATGGAAGAAATGGTCACTATTGCTGCAAGCTC 61829

Qy        728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGl 748
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61830 ACCACTATTTTCACATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGG 61889

Qy        748 yArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61890 CCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCC 61949

Qy        768 oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh 788
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61950 AGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTAC 62009

Qy        788 rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs 808
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62010 TCATGTCATGGGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGA 62069

Qy        808 pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl 828
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62070 CCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGC 62129

Qy        828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62130 AACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCC 62189

Qy        848 oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe 868
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62190 TAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATTAGCAAG 62249

Qy        868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62250 CTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCA 62309

Qy        888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl 908
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62310 GTATGCACATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAATGG 62369

Qy        908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62370 AAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACAGACC 62429

Qy        928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62430 TTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAA 62489

Qy        948 nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleTh 968
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62490 CTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAAC 62549

Qy        968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62550 ATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGC 62609

Qy        988 aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr 1008
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62610 AAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAG 62669

Qy       1008 gAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62670 GAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATCG 62729

Qy       1028 gPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSe 1048
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62730 GTTTGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATTTTACTACAATCAG 62789

Qy       1048 rAsnAspAsnGluTyrMetAspTyrArg 1057
                 |||||||||||||||||||||||||||
```

```
Db      62790 TAACGATAACGAGTATATGGATTACCGC 62817

RESULT 5
AP001859
LOCUS       AP001859               150594 bp    DNA     linear   PLN 19-JAN-2005
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
            PAC clone:P0485D09.
ACCESSION   AP001859 BA000010
VERSION     AP001859.1  GI:7630233
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
            Katayose,Y., Wu,J., Niimura,Y., Cheng,Z., Nagamura,Y.,
            Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
            Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
            Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
            Idonuma,A., Iijima,M., Ikeda,M., Ikeno,M., Ito,S., Ito,T., Ito,Y.,
            Ito,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katagiri,S.,
            Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
            Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
            Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
            Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
            Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
            Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
            Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
            Yano,M., Jiang,J. and Gojobori,T.
  TITLE     The genome sequence and structure of rice chromosome 1
  JOURNAL   Nature 420 (6913), 312-316 (2002)
   PUBMED   12447438
REFERENCE   2  (bases 1 to 150594)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki, 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
            (http://globin.cse.psu.edu/html/docs/sim4.html), gap2
            (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The
            genomic sequence was searched against NCBI NonRedundant Protein
            database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
            sequence database at RGP or DDBJ. Protein homologies of the coding
            regions were searched against NCBI NonRedundant Protein database
            with BLASTP. ESTs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
            Full-length cDNAs represent the identified cDNA sequences using
            BLASTN with the corresponding DDBJ accession no.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'. A gene without
            significant homology to any protein but with full-length cDNA or
            EST homology (covering almost the entire length of partial
            sequence) is classified as an 'unknown' protein. A gene predicted
            by two or more gene prediction programs is classified as a
            'hypothetical' protein according to IRGSP standard. A gene
            predicted by a single gene prediction program is also classified as
            a probable 'hypothetical' protein and is included as a
            miscellaneous feature of the sequence.
            The orientation of the sequence is from T7 to SP6 of the PAC clone.
            This sequence of P0485D09 clone has an overlap with P0452F10(DDBJ:
            AP003434) clone at 5' end and with P0431F01(DDBJ: AP001550) clone
```

```
          at 3' end. Detailed information on overlap and assembly quality
          together with annotation of this entry is available at
          http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source          1. .150594
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="1"
                     /clone="P0485D09"
     gene            complement(4801. .5133)
                     /gene="P0485D09.1"
     mRNA            complement(<4801. .>5133)
                     /gene="P0485D09.1"
                     /note="start and end point are not identified"
     CDS             complement(4801. .5133)
                     /gene="P0485D09.1"
                     /note="predicted by FGENESH etc."
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAD81318.1"
                     /db_xref="GI:56783881"
                     /translation="MTELRSGLGVAGTCAMAHGASCLITLSPQLKDYEDTFSLTGKIQ
                     NFRELLYYSDGLNSEEKRGSRWLGTAANDVSLSDFSPAATTNSPLIMLWSLANLVRLG
                     LVAPWISK"
     gene            join(5701. .5883,6903. .6908)
                     /gene="P0485D09.2"
     misc_feature    join(5701. .5883,6903. .6908)
                     /gene="P0485D09.2"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            join(7919. .7964,8092. .8280,8453. .8473,8558. .8610)
                     /gene="P0485D09.3"
     misc_feature    join(7919. .7964,8092. .8280,8453. .8473,8558. .8610)
                     /gene="P0485D09.3"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            complement(9235. .9582)
                     /gene="P0485D09.4"
     misc_feature    complement(9235. .9582)
                     /gene="P0485D09.4"
                     /note="hypothetical ORF
                     predicted by GENSCAN
                     this category is not included in IRGSP standard"
     gene            10769. .17070
                     /gene="P0485D09.5"
     mRNA            join(<10769. .10884,10933. .10987,11596. .11653,12285. .12379,
                     12471. .12574,12958. .13086,13493. .15281,15497. .15702,
                     16513. .16597,16831. .>17070)
                     /gene="P0485D09.5"
                     /note="start and end point are not identified"
     CDS             join(10769. .10884,10933. .10987,11596. .11653,12285. .12379,
                     12471. .12574,12958. .13086,13493. .15281,15497. .15702,
                     16513. .16597,16831. .17070)
                     /gene="P0485D09.5"
                     /codon_start=1
                     /product="putative Myb-like protein"
                     /protein_id="BAD81319.1"
                     /db_xref="GI:56783882"
                     /translation="MESLLDLRRLLKKEKSAMNQKGDDLSMGKRLKALLTIVMRTTGP
                     TRRSTKGNWTPEEDAILSRAVQTYNGKNWKKIAECFPDRTDVQCLHRWQKVLNPELVK
                     GPWSKEEDEIIVQMVNKLGPKKWSTIAQALPGRIGKQCRERWYNHLNPGINKEAWTQE
                     EEITLIHAHRMYGNKWAELTKFLPGRTDNSIKNHWNSSVKKKVNSYMSSGLLTQVSCL
                     PLNEYSANCNSSPAMTQQNSEDSGCFAVREVENSSGCSQSSLAKVSCSQVHDTTVPLG
                     CDLQVNANFDKNEAHDSQSSMGPQACYTSAEAVASALPAVHCHVSSSNLDPDQHLQED
                     FAQGLNLDMTIDEMPTVPSFADNQTVCSIENHERSLEPYDVAMEVPLSMLSSDSGAEQ
                     KLHFMSEADFNSPNCLKSELWQDISLQGLLSGPDAVEADSISRSNHQSDVYSSEADTH
                     FLAPPYMPQTSNSSSVMGLADDQSPQMSVPPSLICSNAMTDDAPFDNRPGRKEMPLSQ
                     AEVVTQSSSSSGDAEMFANPGCSNDRHVPSSTMESIPECGDQQVTNAEEPEASLEKEP
                     SLTQSVTAPDEQDKGALFYEPPRFPSLDVPFVSCDLVTSGDLQEFSPLGIRQLMHSTM
                     NVCTPMRLWGSPTHDESTGVLLKSAAKSFICTPSILKKRHRDLLSPIPDKRIEKKYGT
                     EKDRGVSDTSSTGIQTSCINATKDDALITTVLRIERSASSKSLEKKLVFSDENKENLG
                     YTTEQTKDGQSAGNDEHMDEQTTGERSSATNVATNDDLSGNLQPAGILIEHSGDDPIS
                     PDYGKNTMNQKLNTNVKSLSVCKEGVCAKSKPTELIVEKSSPCINVDYEYVNILADTP
                     GIKRGLESPSAWKSPWFVDMHFQGSYFTSPADSYDALGLMKQINVQTAAALVEAREVL
                     ASGGQCDNISSDKENTGNPDAKKEPGTTKLQTKIMVRGSELPLR"
     gene            complement(join(19054. .19378,19502. .19573,19905. .19998,
                     20072. .20220,20774. .21113,21407. .21738,21975. .22540))
                     /gene="P0485D09.6-1"
     mRNA            complement(join(19054. .19378,19502. .19573,19905. .19998,
                     20072. .20220,20774. .21113,21407. .21738,21975. .22540))
                     /gene="P0485D09.6-1"
                     /note="supported by full-length cDNA(s): AK070838"
     CDS             complement(join(19334. .19378,19502. .19573,19905. .19998,
                     20072. .20220,20774. .21113,21407. .21738,21975. .22508))
                     /gene="P0485D09.6-1"
                     /note="contains EST(s): AU173357(R2418)
                     contains full-length cDNA(s): AK070838"
                     /codon_start=1
                     /product="putative eukaryotic translation initiation
                     factor 3 subunit 6 interacting protein"
                     /protein_id="BAD81320.1"
                     /db_xref="GI:56783883"
                     /translation="MAAAAAEGGGGGGGAVYDPSYVPDSVKTFVAHMYRHVRDKNVYE
                     IHQMYEGGFQRLSDRLFRDAPWPAAEAVSPYCDGDHVFLLLYRELWYRHAYARASSSS
                     SSSAPLTAGQRAESWANYCDLFSVVLHGVVNMQLPNQWLWDMVDEFVYQFQSFCQYRA
                     KLKNKSDDELHQLKQFDKAWNVYGVLNYLQALVEKSMIAQILEREKEGLEQFTATDGY
                     DYQGGSNVLKMLGYYSMIGLLRIHCLLGDYRTGLKCLAPIDLNQQGVYTIVIGSHISA
                     IYHYGFANLMMRRYAEAIREFNKILLYILKYKQYHQKSPQYDQILKKNEQMYAFLAVC
                     LSLCPQHNLIDENVSTQLKEKYNDKMTKMQRFDEETYAAYDELFSYACPKFITPSPPA
                     LDQPLTNYNQDAYRLQLKLFLYEVKQQQLLSGIRSYLKLYSTITIAKLAQYMEVDEAT
                     LRSILMTYKHKMHAVDNNGKIVSSADFDFYIKEDVIHVMESKPIKRHGDYFLRQILKF
                     EEMIGELEKVQFD"
     gene            complement(join(21153. .21738,21975. .22547))
                     /gene="P0485D09.6-2"
     mRNA            complement(join(21153. .21738,21975. .22547))
                     /gene="P0485D09.6-2"
                     /note="supported by full-length cDNA(s): AK061285"
     CDS             complement(join(21400. .21738,21975. .22508))
                     /gene="P0485D09.6-2"
                     /note="contains full-length cDNA(s): AK061285"
                     /codon_start=1
                     /product="putative eukaryotic translation initiation
                     factor 3 subunit 6 interacting protein"
                     /protein_id="BAD81321.1"
                     /db_xref="GI:56783884"
                     /translation="MAAAAAEGGGGGGGAVYDPSYVPDSVKTFVAHMYRHVRDKNVYE
                     IHQMYEGGFQRLSDRLFRDAPWPAAEAVSPYCDGDHVFLLLYRELWYRHAYARASSSS
                     SSSAPLTAGQRAESWANYCDLFSVVLHGVVNMQLPNQWLWDMVDEFVYQFQSFCQYRA
                     KLKNKSDDELHQLKQFDKAWNVYGVLNYLQALVEKSMIAQILEREKEGLEQFTATDGY
                     DYQGGSNVLKMLGYYSMIGLLRIHCLLGDYRTGLKCLAPIDLNQQGVYTIVIGSHISA
                     IYHYGFANLMMRRY"
     gene            22882. .30975
                     /gene="P0485D09.7"
     mRNA            join(22882. .23145,24592. .24670,25118. .25216,25302. .25415,
                     25608. .25750,26364. .26445,27073. .27158,27527. .27760,
                     29332. .29810,29957. .30975)
                     /gene="P0485D09.7"

Alignment Scores:
Pred. No.:                2.05e-267      Length:          150594
Score:                    5191.00        Matches:         1055
Percent Similarity:       68.1%          Conservative:    0
Best Local Similarity:    68.1%          Mismatches:      2
Query Match:              93.9%          Indels:          495
DB:                       15             Gaps:            3

US-09-828-068-2 (1-1057) x AP001859 (1-150594)

Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      52820 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 52879
```

```
Qy        21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     52880 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 52939

Qy        41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     52940 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 52999

Qy        61 Ile---------------------------------------------------------- 61
Db     53000 AT-AAGGTAATCATTTTCTGTATTTCCAATTCCAGTATCGCGTTGTGGATGAATAATGAA 53058

Qy        61 ---------------------------------------------------------- 61
Db     53059 TCGGCATGTCATGCCATATTGCACTGTTTGATGGAAGAGTATGATTGATCGTGGTTTTTG 53118

Qy        61 ---------------------------------------------------------- 61
Db     53119 CACAGTTTGCTGTTGGGACTTATATGGTCATCTGTTTTGTACGATCGTATACACTGGGTC 53178

Qy        61 ---------------------------------------------------------- 61
Db     53179 GACATGCTTATGACTTTGGTTCGATTTAGGAAGTCAATACATCCACTACTAGCTCTATAT 53238

Qy        61 ---------------------------------------------------------- 61
Db     53239 CTAGCCATGTGAACTCATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAAAAAATATA 53298

Qy        61 ---------------------------------------------------------- 61
Db     53299 TATAATATTTGCATATATGTTGGTGTTTCATGTATCTTTATACTCTACGTACATCCATTA 53358

Qy        61 ---------------------------------------------------------- 61
Db     53359 ATATCTTCAATGTATGAATCTGAGCACATGATTGTGAGTGCTACACATATGCATGTCTGT 53418

Qy        61 ---------------------------------------------------------- 61
Db     53419 ATGTGTGTTCATTAGGTGTTTGATCATATTTGTTTGTGTTGGGGTGCGCATGCATTTATT 53478

Qy        61 ---------------------------------------------------------- 61
Db     53479 CAGGCCATGCTGTAGGCTGTAGCTAGATATTTGTGTTTGTATATTATTTCTGTTGAACAA 53538

Qy        61 ---------------------------------------------------------- 61
Db     53539 GCTGATTACTAATGAAATTAACCTTTTTGGGGTACACTCATATATTGGGCCCTACATTTT 53598

Qy        61 ---------------------------------------------------------- 61
Db     53599 TGTAATCATTTTTCCTTTGTGCTGAGGTTCAGCATAAAACTTTTTTATCATAAGCATGTT 53658

Qy        61 ---------------------------------------------------------- 61
Db     53659 TACATCCTAGGAGATTCTTAGAACTGATGGTTTCTTCATATTTGCATTATGTTTGATTTG 53718

Qy        61 ---------------------------------------------------------- 61
Db     53719 ATAGTCCATTATTATTTTAAGCCTTTTCAATTGTTTAGAGATTCTAGAGATGATATATAT 53778

Qy        61 ---------------------------------------------------------- 61
Db     53779 CAACCATAGACTTGTCACGTTTTGGTTTAATACTTTCTAGAACTAATTAGATTATTATTT 53838

Qy        61 ---------------------------------------------------------- 61
Db     53839 TTGTAGTTTATCCTGTCATGCTATTTGTATTATCTTTGAATTCAAACTGCAATACTTAGA 53898

Qy        61 ---------------------------------------------------------- 61
Db     53899 TTATCTTGAAGGTCCTCTTTTTCTGGACTGTACAAGCTATGTATGAAATGCCTACCTCCC 53958

Qy        61 ---------------------------------------------------------- 61
```

```
Db     53959 AGCATCCTTTAGATTATGTAGGGCCTTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC 54018

Qy        61 ---------------------------------------------------------- 61
Db     54019 ACGCAATGTGCTATATATATGTGCCATGCATCTTTATAATGATAATCTTATTTTTCTTGT 54078

Qy        62 ---ArgGlyTyrValAlaLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54079 ACCAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 54138

Qy        81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54139 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 54198

Qy       101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54199 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 54258

Qy       121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54259 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 54318

Qy       141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54319 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 54378

Qy       161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54379 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 54438

Qy       181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54439 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 54498

Qy       201 AspSerProMetLys---------------------------------------------- 205
             ||||||||||||
Db     54499 GATTCACCAATGAA-AGGTATGGTAGATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTT 54557

Qy       206 ----------------------------------------AspLeuGlnGlyProAla 211
                                                    ||||||||||||||||||
Db     54558 ATTTAAGGTATAGAATAAACTAATGTTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC 54617

Qy       212 GlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAla 231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54618 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 54677

Qy       232 LeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54678 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 54737

Qy       252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54738 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 54797

Qy       272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54798 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGA 54857

Qy       292 LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54858 AAGGAACGTGATCAGGTTGCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCT 54917

Qy       312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54918 GGGCAGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGC 54977

Qy       332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54978 TCCAAATCTAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAA 55037

Qy       352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      55038 CGCACTGCCCAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTG 55097
Qy        372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
Db      55098 CGGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTT 55157
Qy        392 HisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMet 411
Db      55158 CATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATG 55217
Qy        412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
Db      55218 GAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCA 55277
Qy        432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
Db      55278 AGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATG 55337
Qy        452 AsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisPro 471
Db      55338 AACTGGCTGAATGGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCA 55397
Qy        472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
Db      55398 GCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAG 55457
Qy        492 AsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVal 511
Db      55458 AATGATACTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTGTCAGCATGTA 55517
Qy        512 SerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGly 531
Db      55518 TCAGAAATCTCCACACAGAGGTGCTCATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGG 55577
Qy        532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
Db      55578 AAAACACATTCAGCTGCTAGTACCAAATATGGTGGTGAAAGCACCAGAAATGGTCAGAAC 55637
Qy        552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
Db      55638 ATACATGTACTCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGT 55697
Qy        572 HisSerAlaLys---------------------------------------------- 575
Db      55698 CACTCGGCAAAGGTACGAATTTTGTGAATCATGAGGAATTTTTGCTTTTTAAATTGACTG 55757
Qy        575 ---------------------------------------------------------- 575
Db      55758 AATCAACATTTATCTGTATGAAGGAATAATATTGGTGCATAACAATGTTAAGAAATATGC 55817
Qy        575 ---------------------------------------------------------- 575
Db      55818 ATACAATGTTTATTTATATGCTTTCCACTGTTCTTCTTTACTTATGTTTTGATACTCTTT 55877
Qy        575 ---------------------------------------------------------- 575
Db      55878 TTGTGTGTGCGTGCATGTGTGCATGTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGC 55937
Qy        575 ---------------------------------------------------------- 575
Db      55938 ACGTGCGTGGCGCAATATTCTTTTTTAGACTCATATTATAGTGATTGTAATGGACTGACA 55997
Qy        576 ----------------------ValSerProAlaGluHisAspIleGlnIleMetSerAs 588
Db      55998 TTTTCCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGA 56057
Qy        588 pLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGl 608
Db      56058 CCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGA 56117
Qy        608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGl 628
Db      56118 AAAACAGACCATGATAGATGACATCCCCATGGATATTGTTGAACTGCTAGCTAAAAACCA 56177

Qy        628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLy 648
Db      56178 GCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCTGACATCAACCGTATTCAATCCAA 56237
Qy        648 sThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe 668
Db      56238 GACAACTGCTGATGATGATTGTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCATC 56297
Qy        668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGl 688
Db      56298 AAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGA 56357
Qy        688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSe 708
Db      56358 GTTACAGGGTCATTTGGCATTGACCACACAAGAGTCTCCACATCCTCAGAACTTTCAGTC 56417
Qy        708 rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
Db      56418 TACTCAGGAACAGCAGACACATTTGCGGATGGAAGAAATGGTCACTATTGCTGCAAGCTC 56477
Qy        728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGl 748
Db      56478 ACCACTATTTTCACATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGG 56537
Qy        748 yArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
Db      56538 CCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCC 56597
Qy        768 oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh 788
Db      56598 AGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTAC 56657
Qy        788 rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs 808
Db      56658 TCATGTCATGGGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGA 56717
Qy        808 pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl 828
Db      56718 CCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGC 56777
Qy        828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
Db      56778 AACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCC 56837
Qy        848 oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe 868
Db      56838 TAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATTAGCAAG 56897
Qy        868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888
Db      56898 CTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCA 56957
Qy        888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl 908
Db      56958 GTATGCACATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAATGG 57017
Qy        908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
Db      57018 AAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACAGACC 57077
Qy        928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948
Db      57078 TTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAA 57137
Qy        948 nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleTh 968
Db      57138 CTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAAC 57197
Qy        968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
Db      57198 ATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGC 57257
```

```
Qy        988 aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57258 AAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAG 57317

Qy       1008 gAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57318 GAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATCG 57377

Qy       1028 gPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSe 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57378 GTTTGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATTTTACTACAATCAG 57437

Qy       1048 rAsnAspAsnGluTyrMetAspTyrArg 1057
              |||||||||||||||||||||||||||
Db      57438 TAACGATAACGAGTATATGGATTACCGC 57465
```

RESULT 6
AK103364

LOCUS	AK103364 3626 bp mRNA linear PLN 24-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J033126N23, full insert sequence.
ACCESSION	AK103364
VERSION	AK103364.1 GI:32988573
KEYWORDS	FLI_CDNA; CAP trapper.
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
PUBMED	12869764
REFERENCE	2 (bases 1 to 3626)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE	Direct Submission
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cDNA/ NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M. FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES	Location/Qualifiers
source	1. .3626
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="mRNA"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/clone="J033126N23"

ORIGIN

Alignment Scores:

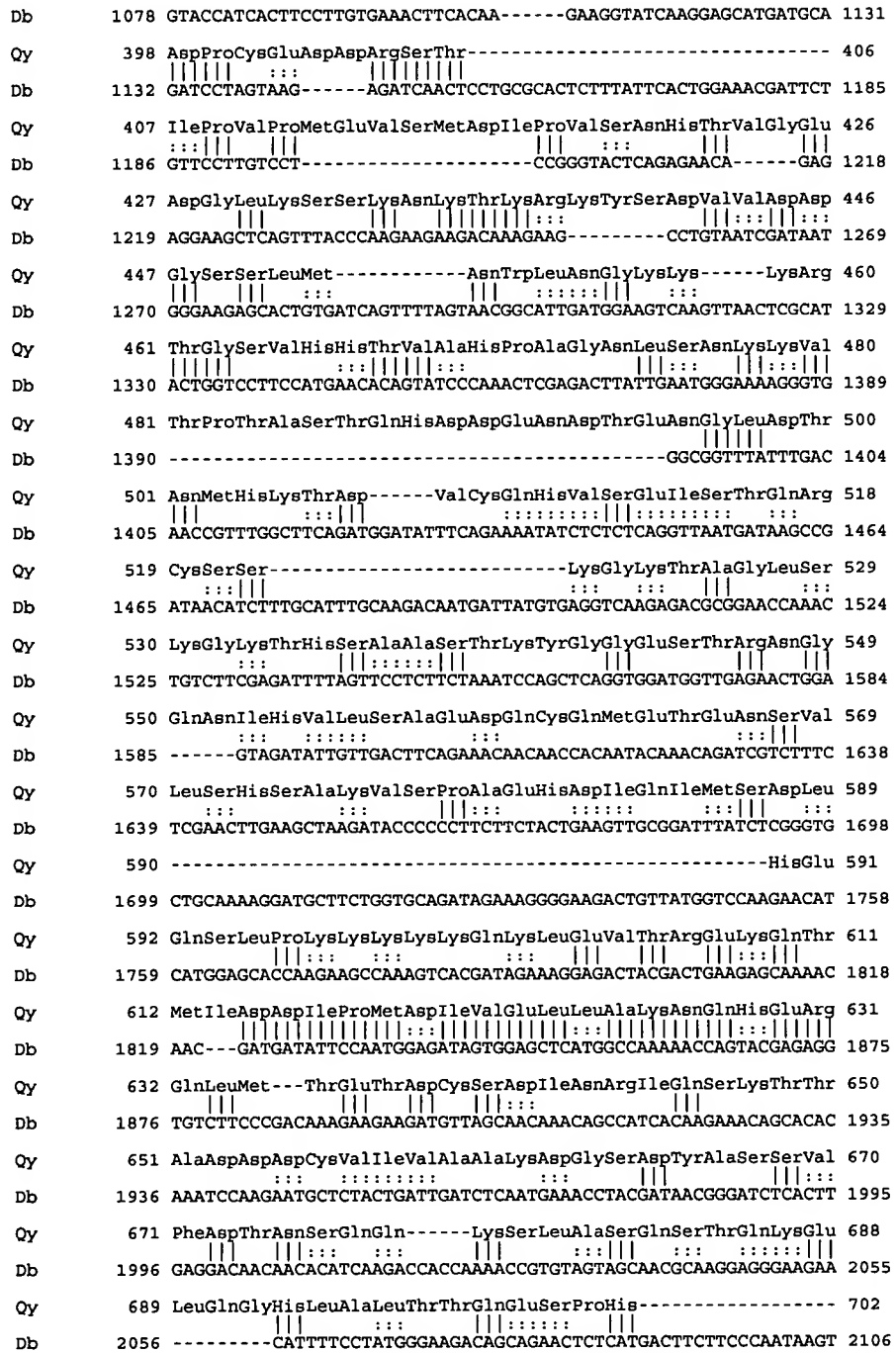
Pred. No.:	1.12e-268	Length:	3626
Score:	5179.00	Matches:	992
Percent Similarity:	99.9%	Conservative:	0
Best Local Similarity:	99.9%	Mismatches:	1
Query Match:	93.7%	Indels:	0
DB:	15	Gaps:	0

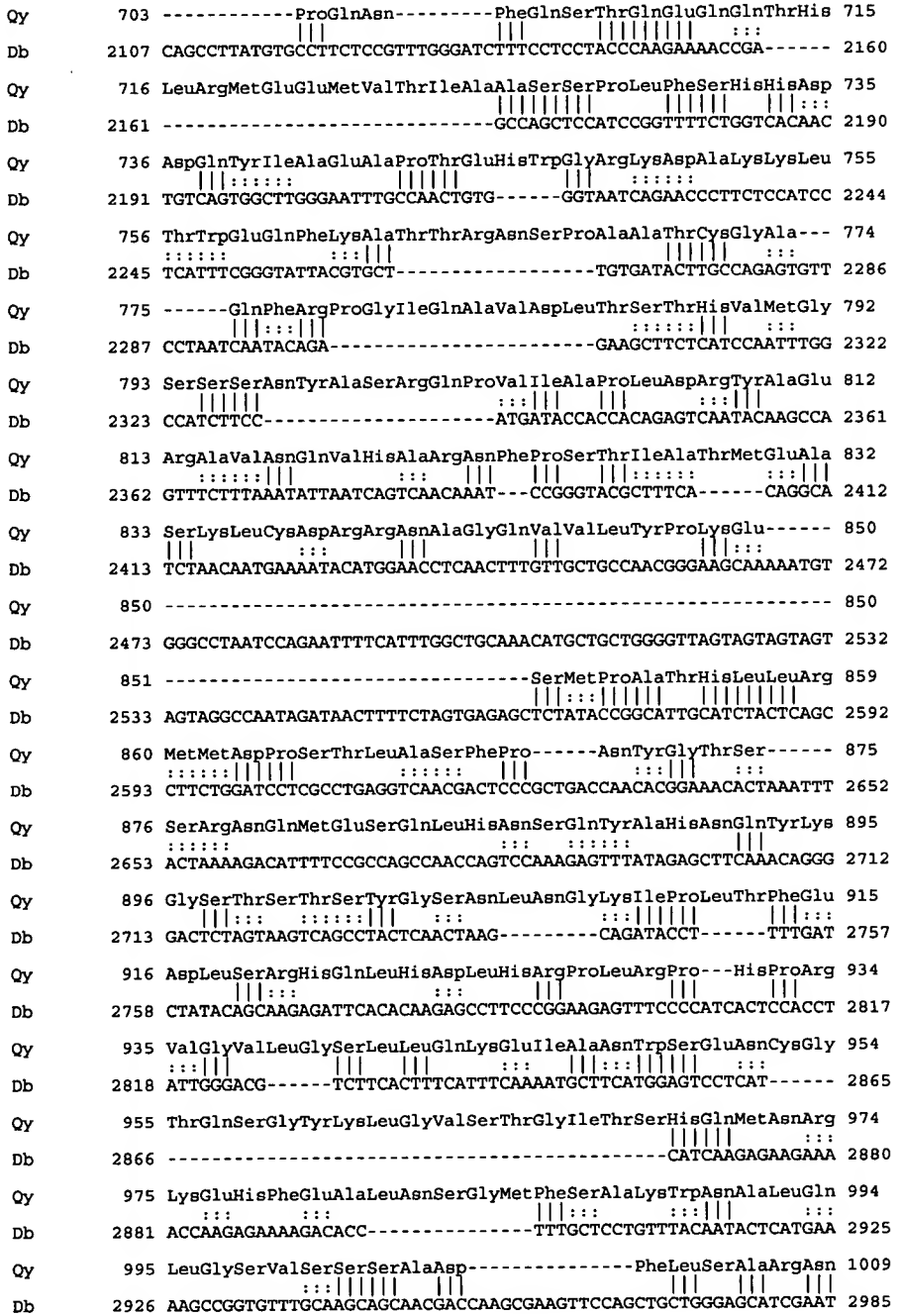
US-09-828-068-2 (1-1057) x AK103364 (1-3626)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 706

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 766

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 826

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 886

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 946
```

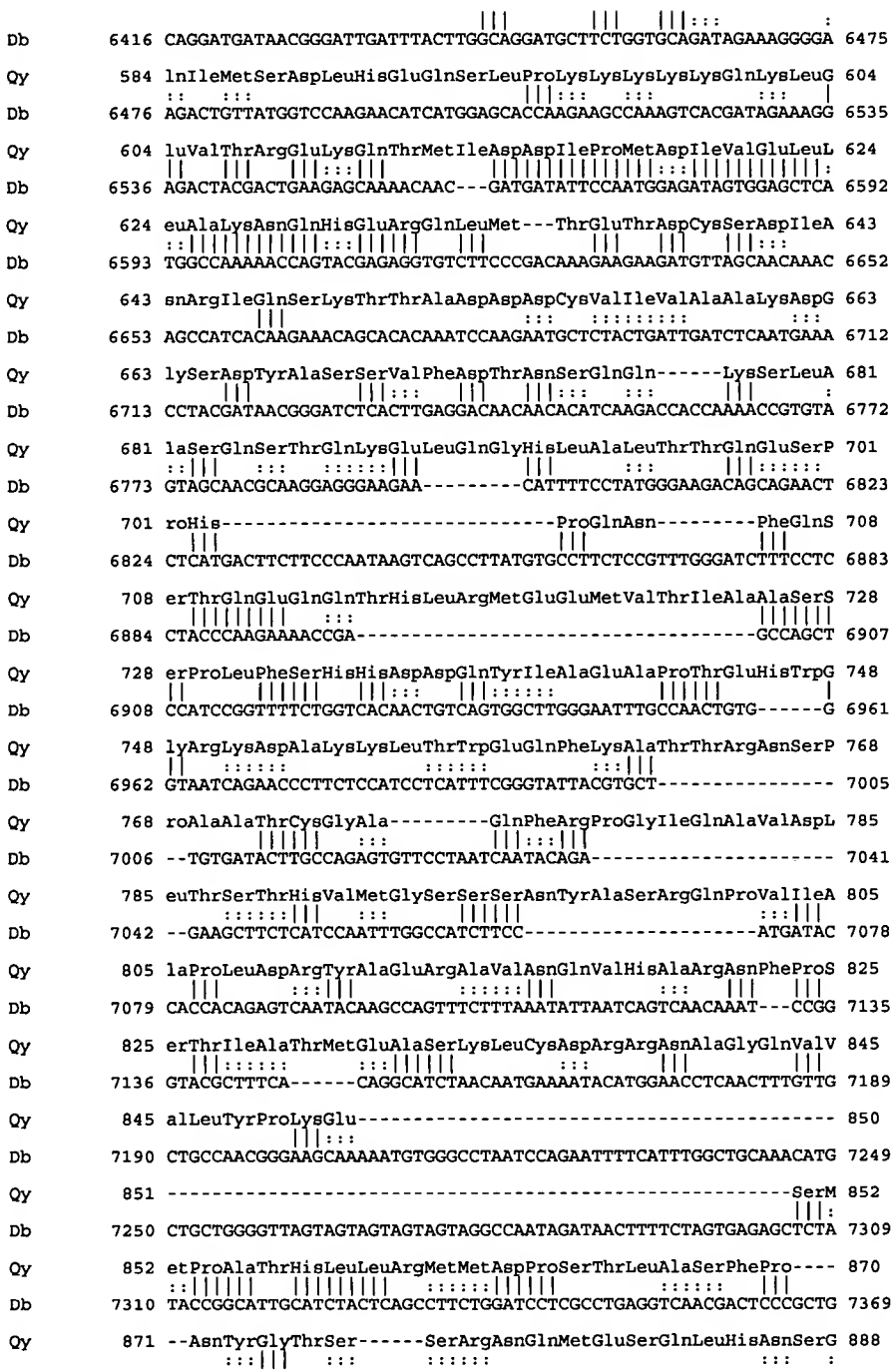
```
Qy      101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
Db      947 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1006

Qy      121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
Db      1007 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1066

Qy      141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
Db      1067 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1126

Qy      161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
Db      1127 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1186

Qy      181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
Db      1187 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1246

Qy      201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
Db      1247 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1306

Qy      221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
Db      1307 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1366

Qy      241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
Db      1367 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1426

Qy      261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
Db      1427 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1486

Qy      281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
Db      1487 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1546

Qy      301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
Db      1547 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1606

Qy      321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
Db      1607 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1666

Qy      341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
Db      1667 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1726

Qy      361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
Db      1727 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1786

Qy      381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
Db      1787 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1846

Qy      401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
Db      1847 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1906

Qy      421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db      1907 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1966

Qy      441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
Db      1967 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2026

Qy      461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480

Db      2027 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2086

Qy      481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
Db      2087 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2146

Qy      501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
Db      2147 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2206

Qy      521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
Db      2207 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2266

Qy      541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
Db      2267 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2326

Qy      561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
Db      2327 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2386

Qy      581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
Db      2387 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2446

Qy      601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db      2447 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2506

Qy      621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db      2507 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2566

Qy      641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db      2567 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2626

Qy      661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db      2627 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2686

Qy      681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db      2687 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2746

Qy      701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
Db      2747 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2806

Qy      721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
Db      2807 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2866

Qy      741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
Db      2867 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2926

Qy      761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
Db      2927 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATATGGTGCTCAATTTAGACCTGGTATC 2986

Qy      781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
Db      2987 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3046

Qy      801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
Db      3047 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3106

Qy      821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
```

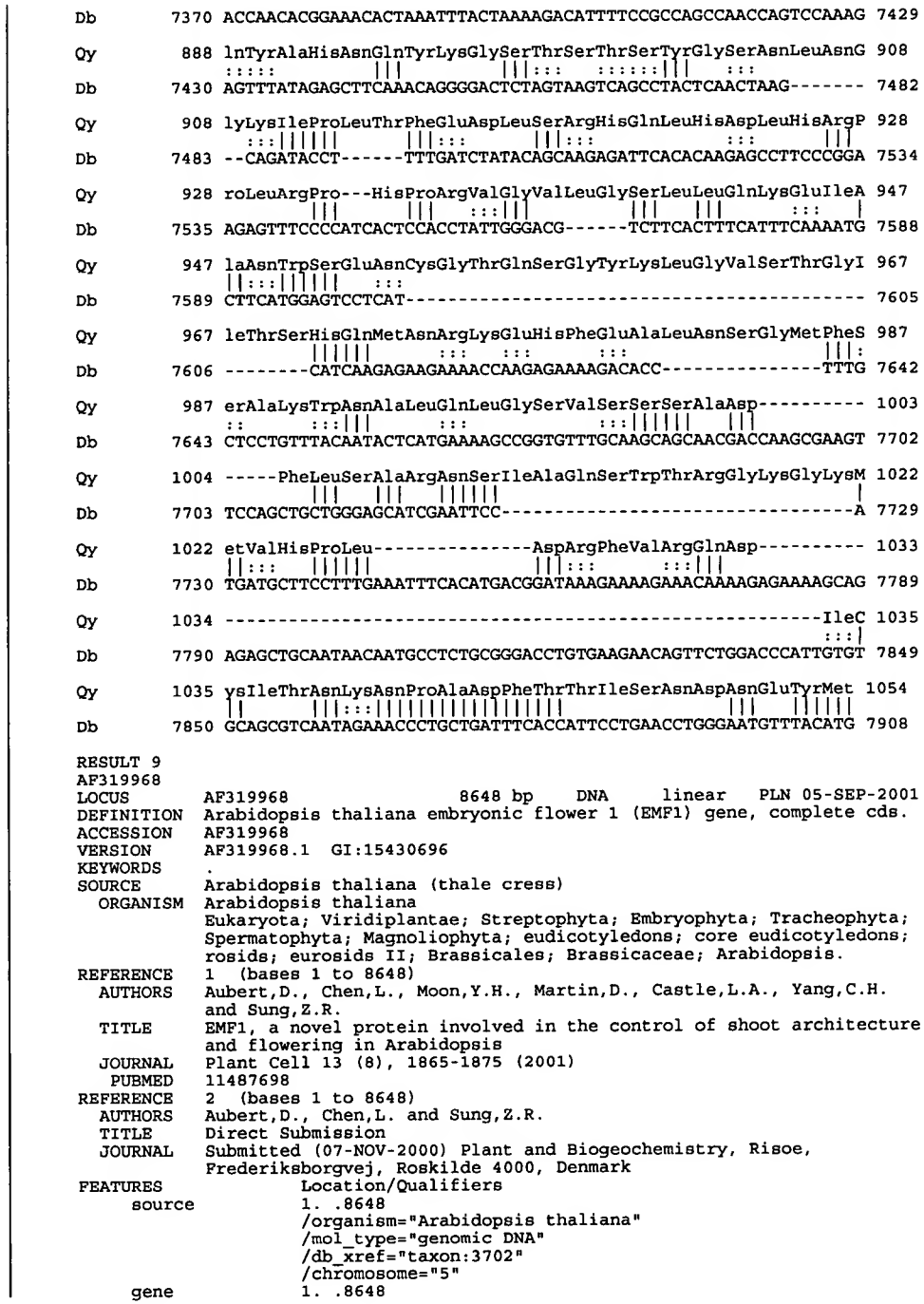
```
Db       3107 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3166

Qy        841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3167 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3226

Qy        861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3227 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3286

Qy        881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3287 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3346

Qy        901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3347 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3406

Qy        921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3407 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3466

Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3467 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3526

Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3527 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3586

Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeu 993
              |||||||||||||||||||||||||||||||||||||||
Db       3587 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTG 3625

RESULT 7
AX590041
LOCUS       AX590041               3291 bp    DNA     linear   PAT 24-JAN-2003
DEFINITION  Sequence 223 from Patent WO02081695.
ACCESSION   AX590041
VERSION     AX590041.1  GI:27901181
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Zhu,T., Glazov,E.A., Meins,F., Wang,X. and Chang,H.S.
  TITLE     Genes that are modulated by posttranscriptional gene silencing
  JOURNAL   Patent: WO 02081695-A 223 17-OCT-2002;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1. .3291
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN

Alignment Scores:
Pred. No.:              6.79e-12        Length:         3291
Score:                  393.00          Matches:        248
Percent Similarity:     36.8%           Conservative:   191
Best Local Similarity:  20.8%           Mismatches:     401
Query Match:            7.1%            Indels:         352
DB:                     6               Gaps:           58

US-09-828-068-2 (1-1057) x AX590041 (1-3291)

Qy          55 GlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLys 74
               :::|||:::|||||||||:::|||||||:::||||||      :::::::::|||   :::
Db          73 AAATGTGATCATTTTTCCATGCGTGGATTCGTAGCTGAAACTCGTGAGAGAGACCTTAGA 132

Qy          75 PheCys-----------------SerLeuSerArgIlePheHisAspGlnLysLysCys 88
                  |||                   |||:::|||   :::
Db         133 AAATGTTGGCCGTTTTCAGAAGAGAGTGTTAGTTTAGTA-------------------- 171

Qy          89 AspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCys 108
               |||:::         :::              ||||||   |||||||||   |||   |||
Db         172 GACCAACAAAGCTATACTCTTCCTACTTTATCTGTTCCAAAGTTTAGATGGTGGCATTGT 231

Qy         109 SerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
                  |||:::      :::            :::          |||:::
Db         232 ATGAGCTGCATCAAAGATATA-----------GATGCTCATGGGCCAAAAGATTGTGGA 279

Qy         129 AlaLysGlnAsnGlyThrSerAspGly---CysSerIleThrPheValArgSerThrPhe 147
                  |||      :::   |||      |||:::          :::|||   |||
Db         280 CTGCATTCAAACTCAAAAGCTATTGGAAACTCTTCTGTTATCGAAAGTAAAAGCAAGTTC 339

Qy         148 ValProAlaSerValGlySerGlnLysValSerProSerThr----------------- 161
                  ::::::       :::              |||
Db         340 AATTCGCTAACTATCATTGATCACGAGAAAGAAAAGAAAACTGATATTGCAGATAATGCT 399

Qy         162 ---GlnSerSerGlnGlyLysAsnAlaAspArg--------------SerThrLeuPro 175
                  :::         |||   |||   :::                    :::|||
Db         400 ATTGAGGAGAAAGTGGGTGTAAACTGTGAGAATGATGATCAGACAGCTACTACGTTTCTC 459

Qy         176 LysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAla 195
               |||      :::                          |||   |||
Db         460 AAGAAAGCACGT--------------------GGTCGACCTATGGGT------------ 486

Qy         196 AlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAsp 215
                  :::|||      |||      :::   |||      |||   |||
Db         487 ---GCTTCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG--------- 534

Qy         216 ValAlaAlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeu 232
               |||      |||   |||::::::               |||:::|||:::   :::
Db         535 GTAGGAAACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGG 594

Qy         233 ProGluValProGlnIleThrTrpHisIleGluValAsnGlyAla--------------- 247
                  |||         :::         :::         |||:::
Db         595 AAAGAGAAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGT 654

Qy         248 ---------AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsn 264
                     |||   ||||||         |||      :::   :::
Db         655 GTGGTTGAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATG 714

Qy         265 GluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLys 284
               |||   :::||||||   :::|||          :::|||
Db         715 GAATGCGATAACGGGTCATCAGAAAGTATAAATCTTGCT---------------------- 753

Qy         285 AspProAsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCys 301
                        ||||||||   :::|||              :::|||
Db         754 ------------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGT 801

Qy         302 AsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGlu 321
                  |||      :::   |||      |||                 :::
Db         802 GAGTTGCTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA------------ 849

Qy         322 ProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLys 341
                  ||||||      ||||||:::   |||          ||||||      :::|||
Db         850 ------GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAG 897

Qy         342 LeuMetLysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp 360
               |||:::   :::      :::   |||:::      :::                |||:::
Db         898 TTGTTACCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAA 957

Qy         361 ------AlaLysLeuCys------------------------------------------ 364
                     :::|||   |||
Db         958 AATGCTTCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTT 1017

Qy         365 --------------------ArgArgLysProLysLysValArgLeuLeuSerGluIle 377
                                   :::::::::   ::::::   :::::::::   |||
Db        1018 GACAGAACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTT 1077

Qy         378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397
               :::   :::          |||   |||:::      |||      :::|||:::   |||
```





```
Qy       1010 SerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeu--------- 1026
Db       2986 TCC--------------------------------ATGATGCTTCCTTTGAAATTTCAC 3012

Qy       1027 ------AspArgPheValArgGlnAsp------------------------------- 1033
Db       3013 ATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAGAGAGCTGCAATAACAATGCCTCT 3072

Qy       1034 --------------------------------IleCysIleThrAsnLysAsnProAla 1042
Db       3073 GCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGTGCAGCGTCAATAGAAACCCTGCT 3132

Qy       1043 AspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db       3133 GATTTCACCATTCCTGAACCTGGGAATGTTTACATG 3168

RESULT 8
AR368885
LOCUS       AR368885               8648 bp    DNA     linear   PAT 12-SEP-2003
DEFINITION  Sequence 2 from patent US 6376751.
ACCESSION   AR368885
VERSION     AR368885.1  GI:34603269
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 8648)
  AUTHORS   Sung,Z.R., Aubert,D. and Chen,L.
  TITLE     Nucleic acids encoding EMF1 that control reproductive development
            in plants
  JOURNAL   Patent: US 6376751-A 2 23-APR-2002;
            The Regents of the University of California; Oakland, CA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .8648
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.65e-06        Length:         8648
Score:                  298.00          Matches:        225
Percent Similarity:     35.5%           Conservative:   152
Best Local Similarity:  21.2%           Mismatches:     360
Query Match:            5.4%            Indels:         324
DB:                     6               Gaps:           50

US-09-828-068-2 (1-1057) x AR368885 (1-8648)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db       5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
Db       5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
Db       5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Db       5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy        267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db       5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT------------------------- 5353

Qy        287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
Db       5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy        304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db       5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db       5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy        344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
Db       5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy        361 AlaLysLeuCys------------------------------------------------ 364
Db       5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy        365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db       5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy        380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db       5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy        400 CysGluAspAspArgSerThr--------------------------------IlePro 408
Db       5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy        409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db       5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy        429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db       5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy        449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
Db       5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db       5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
Db       5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy        500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db       6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy        516 ThrGlnArgCys------------------SerSerLysGlyLysThrAlaGly---Leu 528
Db       6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy        529 SerLysGly--------------------------------------------LysThr 533
Db       6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy        534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db       6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy        554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db       6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy        573 r------------------------------------------------------------ 573
Db       6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy        574 ----------------------------Ala-LysValSerProAlaGluHisAspIleG 584
```





RESULT 9
AF319968
LOCUS AF319968 8648 bp DNA linear PLN 05-SEP-2001
DEFINITION Arabidopsis thaliana embryonic flower 1 (EMF1) gene, complete cds.
ACCESSION AF319968
VERSION AF319968.1 GI:15430696
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 8648)
AUTHORS Aubert,D., Chen,L., Moon,Y.H., Martin,D., Castle,L.A., Yang,C.H. and Sung,Z.R.
TITLE EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis
JOURNAL Plant Cell 13 (8), 1865-1875 (2001)
PUBMED 11487698
REFERENCE 2 (bases 1 to 8648)
AUTHORS Aubert,D., Chen,L. and Sung,Z.R.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2000) Plant and Biogeochemistry, Risoe, Frederiksborgvej, Roskilde 4000, Denmark
FEATURES Location/Qualifiers
source 1. .8648
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
gene 1. .8648

```
                     /gene="EMF1"
     promoter        1. .3201
                     /gene="EMF1"
     mRNA            join(3202. .3265,4160. .4335,4448. .4623,4704. .4823,
                     4903. .4956,5046. .6307,6448. .8648)
                     /gene="EMF1"
                     /product="embryonic flower 1"
     CDS             join(4241. .4335,4448. .4623,4704. .4823,4903. .4956,
                     5046. .6307,6448. .8031)
                     /gene="EMF1"
                     /codon_start=1
                     /product="embryonic flower 1"
                     /protein_id="AAK98528.1"
                     /db_xref="GI:15430697"
                     /translation="MGSSIKINSISIDLAGAANEIDMVKCDHFSMRGFVAETRERDLR
                     KCWPFSEESVSLVDQQSYTLPTLSVPKFRWWHCMSCIKDIDAHGPKDCGLHSNSKAIG
                     NSSVIESKSKFNSLTIIDHEKEKKTDIADNAIEEKVGVNCENDDQTATTFLKKARGRP
                     MGASNVRSKSRKLVSPEQVGNNRSKEKLNKPSMDISSWKEKQNVDQAVTTFGSSEIAG
                     VVEDTPPKATKNHKGIRGLMECDNGSSESINLAMSGLQRRKSRKVRLLSELLGNTKTS
                     GGSNIRKEESALKKESVRGRKRKLLPENNYVSRILSTMGATSENASKSCDSDQGNSES
                     TDSGFDRTPFKGKQRNRRFQVVDEFVPSLPCETSQEGIKEHDADPSKRSTPAHSLFTG
                     NDSVPCPPGTQRTERKLSLPKKKTKKPVIDNGKSTVISFSNGIDGSQVNSHTGPSMNT
                     VSQTRDLLNGKRVGGLFDNRLASDGYFRKYLSQVNDKPITSLHLQDNDYVRSRDAEPN
                     CLRDFSSSSKSSSGGWLRTGVDIVDFRNNNHNTNRSSFSNLKLRYPPSSTEVADLSRV
                     LQKDASGADRKGKTVMVQEHHGAPRSQSHDRKETTTEEQNNDDIPMEIVELMAKNQYE
                     RCLPDKEEDVSNKQPSQETAHKSKNALLIDLNETYDNGISLEDNNTSRPPKPCSSNAR
                     REEHFPMGRQQNSHDFFPISQPYVPSPFGIFPPTQENRASSIRFSGHNCQWLGNLPTV
                     GNQNPSPSSFRVLRACDTCQSVPNQYREASHPIWPSSMIPPQSQYKPVSLNINQSTNP
                     GTLSQASNNENTWNLNFVAANGKQKCGPNPEFSFGCKHAAGVSSSSSRPIDNFSSESS
                     IPALHLLSLLDPRLRSTTPADQHGNTKFTKRHFPPANQSKEFIELQTGDSSKSAYSTK
                     QIPFDLYSKRFTQEPSRKSFPITPPIGTSSLSFQNASWSPHHQEKKTKRKDTFAPVYN
                     THEKPVFASSNDQAKFQLLGASNSMMLPLKFHMTDKEKKQKRKAESCNNNASAGPVKN
                     SSGPIVCSVNRNPADFTIPEPGNVYMLTGEHLKVRKRTTFKKKPAVCKQDAMKQTKKP
                     VCPPTQNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.65e-06        Length:          8648
Score:                    298.00          Matches:         225
Percent Similarity:       35.5%           Conservative:    152
Best Local Similarity:    21.2%           Mismatches:      360
Query Match:              5.4%            Indels:          324
DB:                       15              Gaps:            50

US-09-828-068-2 (1-1057) x AF319968 (1-8648)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
              :::|||      |||      :::   |||      |||   |||          |||
Db       5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
                 |||   |||::::::           |||:::|||:::   :::       |||
Db       5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
                   :::         :::         |||:::
Db       5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
                 |||   ||||||          |||      :::   :::           |||
Db       5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy        267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
              :::||||||   :::|||         :::|||
Db       5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT------------------------- 5353

Qy        287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
                  |||||||||   :::|||           :::|||                |||
Db       5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy        304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
                  :::   |||          |||          :::
Db       5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
              ||||||      ||||||:::   |||         ||||||      :::||||||:::
Db       5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy        344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
                  :::      :::   |||:::      :::           |||:::
Db       5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy        361 AlaLysLeuCys---------------------------------------------- 364
              :::|||   |||
Db       5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy        365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
                           :::::::::   ::::::   :::::::::   |||   :::
Db       5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy        380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
              :::      |||   |||:::      |||      :::|||:::   |||||||||
Db       5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy        400 CysGluAspAspArgSerThr--------------------------------IlePro 408
                 :::      |||||||||                                :::|||
Db       5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy        409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
                 |||                  |||   :::      |||      |||
Db       5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy        429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
              |||      |||   |||||||||:::           |||:::|||:::|||
Db       5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy        449 SerLeuMet-----------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
              |||   :::           |||   ::::::|||   :::             ||||||
Db       5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
                    :::||||||:::               |||:::   |||:::|||
Db       5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
                       :::      :::           |||
Db       5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy        500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
                          |||:::::|||      |||           |||      :::
Db       6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy        516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
                    |||                  |||||||||      :::   |||   |||
Db       6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy        529 SerLysGly-------------------------------------------LysThr 533
                  |||                                            :::
Db       6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy        534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
                 |||           :::|||      ||||||
Db       6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy        554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
              ||||||                |||         :::      ::::::||||||||||
Db       6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy        573 r------------------------------------------------------------ 573
              |
Db       6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy        574 --------------------------Ala-LysValSerProAlaGluHisAspIleG 584
                                       |||   |||:::          :
Db       6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475
```

```
Qy       584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG 604
Db      6476 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6535

Qy       604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db      6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592

Qy       624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db      6593 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 6652

Qy       643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db      6653 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 6712

Qy       663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
Db      6713 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 6772

Qy       681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db      6773 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 6823

Qy       701 roHis----------------------------ProGlnAsn---------PheGlnS 708
Db      6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 6883

Qy       708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db      6884 CTACCCAAGAAAACCGA----------------------------------GCCAGCT 6907

Qy       728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db      6908 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 6961

Qy       748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db      6962 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT---------------- 7005

Qy       768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
Db      7006 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------------------- 7041

Qy       785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Db      7042 --GAAGCTTCTCATCCAATTTGGCCATCTTCC--------------------ATGATAC 7078

Qy       805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db      7079 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135

Qy       825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db      7136 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 7189

Qy       845 alLeuTyrProLysGlu---------------------------------------- 850
Db      7190 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 7249

Qy       851 ----------------------------------------------------SerM 852
Db      7250 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309

Qy       852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
Db      7310 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 7369

Qy       871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db      7370 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 7429

Qy       888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908

Db      7430 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 7482

Qy       908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db      7483 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 7534

Qy       928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db      7535 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 7588

Qy       947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db      7589 CTTCATGGAGTCCTCAT---------------------------------------- 7605

Qy       967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db      7606 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------TTTG 7642

Qy       987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
Db      7643 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 7702

Qy      1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db      7703 TCCAGCTGCTGGGAGCATCGAATTCC--------------------------------A 7729

Qy      1022 etValHisProLeu---------------AspArgPheValArgGlnAsp---------- 1033
Db      7730 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 7789

Qy      1034 ------------------------------------------------------IleC 1035
Db      7790 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 7849

Qy      1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db      7850 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 7908
```

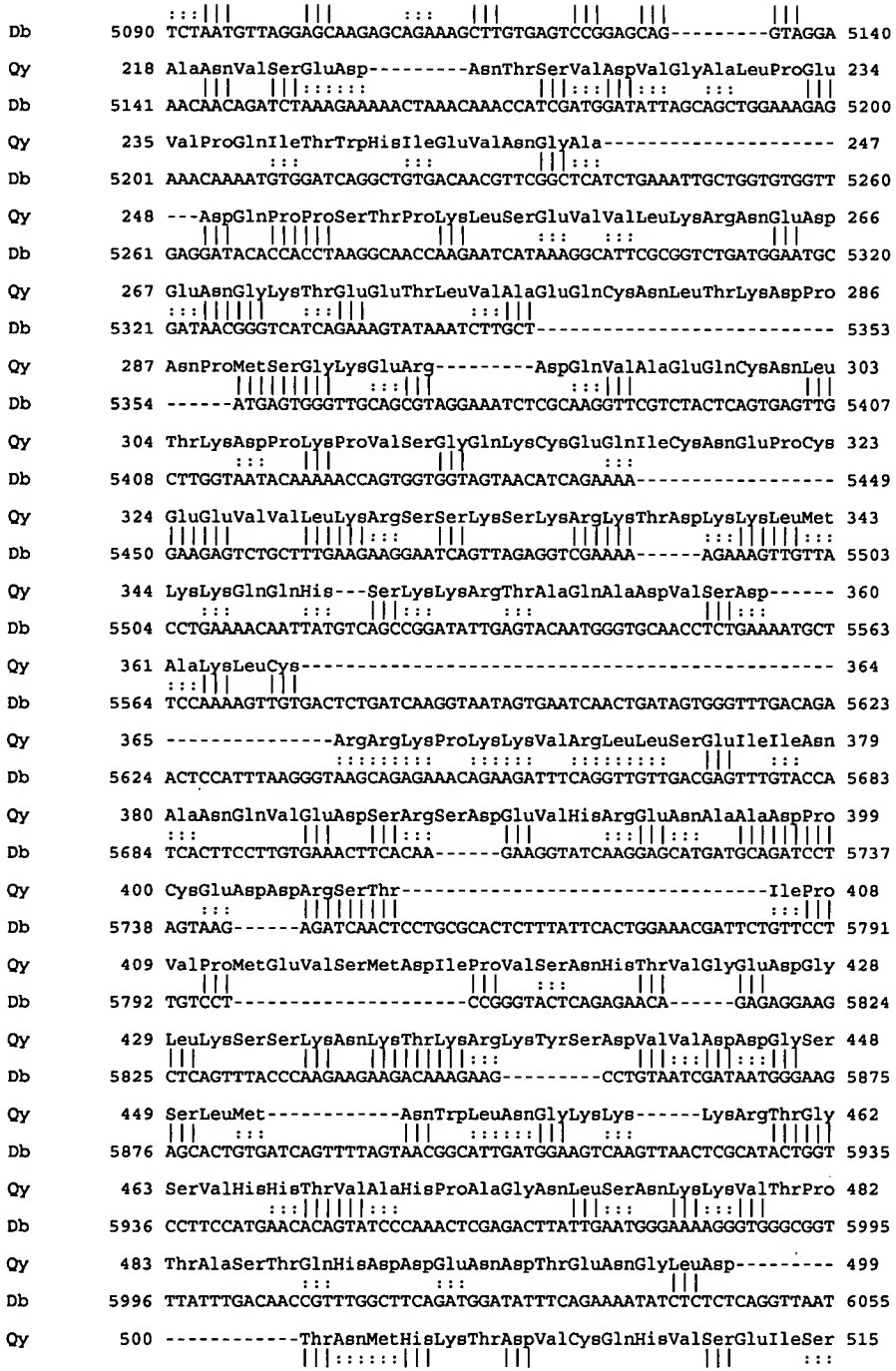
RESULT 10
AR368884
LOCUS AR368884 17341 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 1 from patent US 6376751.
ACCESSION AR368884
VERSION AR368884.1 GI:34603268
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 17341)
AUTHORS Sung,Z.R., Aubert,D. and Chen,L.
TITLE Nucleic acids encoding EMF1 that control reproductive development in plants
JOURNAL Patent: US 6376751-A 1 23-APR-2002;
The Regents of the University of California; Oakland, CA;
WOX;
FEATURES Location/Qualifiers
source 1. .17341
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:

Pred. No.:	6.01e-06	Length:	17341
Score:	298.00	Matches:	225
Percent Similarity:	35.5%	Conservative:	152
Best Local Similarity:	21.2%	Mismatches:	360
Query Match:	5.4%	Indels:	324
DB:	6	Gaps:	50

US-09-828-068-2 (1-1057) x AR368884 (1-17341)

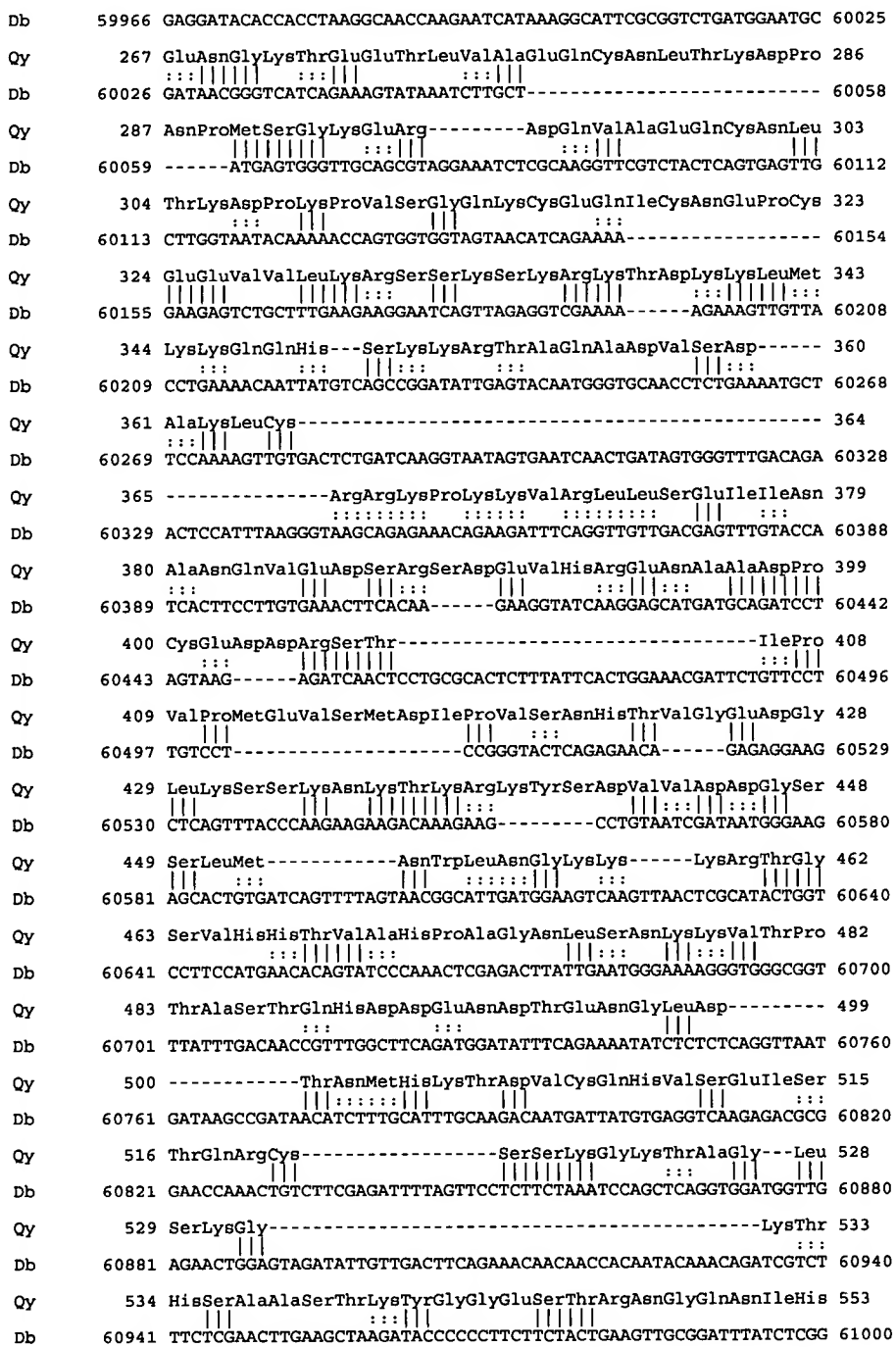
```
Qy       198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
```


Db 5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140
Qy 218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
Db 5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200
Qy 235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------- 247
Db 5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260
Qy 248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Db 5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320
Qy 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db 5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT---------- 5353
Qy 287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
Db 5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407
Qy 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db 5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA---------- 5449
Qy 324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db 5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503
Qy 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
Db 5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563
Qy 361 AlaLysLeuCys---------- 364
Db 5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623
Qy 365 ---------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db 5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683
Qy 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db 5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737
Qy 400 CysGluAspAspArgSerThr----------IlePro 408
Db 5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791
Qy 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db 5792 TGTCCT----------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824
Qy 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db 5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875
Qy 449 SerLeuMet----------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
Db 5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935
Qy 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db 5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
Db 5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055
Qy 500 ----------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db 6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115
Qy 516 ThrGlnArgCys----------SerSerLysGlyLysThrAlaGly---Leu 528
Db 6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175
Qy 529 SerLysGly----------LysThr 533
Db 6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235
Qy 534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db 6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295
Qy 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db 6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355
Qy 573 r---------- 573
Db 6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415
Qy 574 ----------Ala-LysValSerProAlaGluHisAspIleG 584
Db 6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475
Qy 584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Db 6476 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6535
Qy 604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db 6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592
Qy 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db 6593 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 6652
Qy 643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db 6653 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 6712
Qy 663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
Db 6713 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 6772
Qy 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db 6773 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 6823
Qy 701 roHis----------ProGlnAsn---------PheGlnS 708
Db 6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 6883
Qy 708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db 6884 CTACCCAAGAAAACCGA----------GCCAGCT 6907
Qy 728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db 6908 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 6961
Qy 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 6962 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT---------- 7005
Qy 768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
Db 7006 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------- 7041
Qy 785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Db 7042 --GAAGCTTCTCATCCAATTTGGCCATCTTCC----------ATGATAC 7078

```
Qy        805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db       7079 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135

Qy        825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db       7136 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 7189

Qy        845 alLeuTyrProLysGlu---------------------------------------------- 850
Db       7190 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 7249

Qy        851 ------------------------------------------------------SerM 852
Db       7250 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309

Qy        852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
Db       7310 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 7369

Qy        871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db       7370 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 7429

Qy        888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
Db       7430 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 7482

Qy        908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db       7483 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 7534

Qy        928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db       7535 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 7588

Qy        947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db       7589 CTTCATGGAGTCCTCAT---------------------------------------------- 7605

Qy        967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db       7606 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------TTTG 7642

Qy        987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
Db       7643 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 7702

Qy       1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db       7703 TCCAGCTGCTGGGAGCATCGAATTCC-------------------------------A 7729

Qy       1022 etValHisProLeu---------------AspArgPheValArgGlnAsp---------- 1033
Db       7730 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 7789

Qy       1034 ------------------------------------------------------IleC 1035
Db       7790 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 7849

Qy       1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db       7850 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 7908
```

```
RESULT 11
ATF15N18
LOCUS       ATF15N18              101665 bp    DNA     linear   PLN 16-APR-2005
DEFINITION  Arabidopsis thaliana DNA chromosome 5, BAC clone F15N18 (ESSA
            project).
ACCESSION   AL163815
VERSION     AL163815.1  GI:7573398
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 101665)
  AUTHORS   Bevan,M., Hilbert,H., Braun,M., Holzer,E., Brandt,A.,
            Duesterhoeft,A., Bancroft,I., Mewes,H.W., Rudd,S., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 101665)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .101665
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /ecotype="Columbia"
     gene            6461. .8797
                     /gene="F15N18_10"
     CDS             join(6461. .6536,7295. .7792,8271. .8797)
                     /gene="F15N18_10"
                     /note="similarity to predicted proteins in castor bean,
                     Arabidopsis thaliana and alfalfa."
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB87702.1"
                     /db_xref="GI:7573399"
                     /db_xref="GOA:Q9LYE7"
                     /db_xref="InterPro:IPR006946"
                     /db_xref="UniProt/TrEMBL:Q9LYE7"
                     /translation="MKGGSLSFLFVLLIATITSVICFSDGMLPNGDFELGPKPSDMKG
                     TQVINKKAIPSWELSGFVEYIKSGQKQGDMLLVVPAGKFAIRLGNEASIKQRLNVTKG
                     MYYSLTFSAARTCAQDERLNISVAPDSGVIPIQTVYSSSGWDLYAWAFQAESNVAEIV
                     IHNPGEEEDPACGPLIDGVAIKALYPPRPTNKNILKNGGFEEGPYVLPNATTGVLVPP
                     FIEDDHSPLPAWMVESLKAIKYVDVEHFSVPQGRRAVELVAGKESAIAQVARTVVGKT
                     YVLSFAVGDANNACQGSMVVEAFAGKDTLKVPYESRGKGGFKRASLRFVAVSTRTRVM
                     FYSTFYSMRSDDFSSLCGPVIDDVKLLSARKP"
     exon            6461. .6536
                     /gene="F15N18_10"
                     /number=1
     intron          6537. .7294
                     /gene="F15N18_10"
                     /number=1
     exon            7295. .7792
                     /gene="F15N18_10"
                     /number=2
     intron          7793. .8270
                     /gene="F15N18_10"
                     /number=2
     exon            8271. .8797
                     /gene="F15N18_10"
                     /number=3
     gene            10281. .14062
                     /gene="F15N18_20"
     CDS             join(10281. .11965,12429. .12549,13145. .13204,13313. .14062)
                     /gene="F15N18_20"
                     /note="similarity to putative transcription elongation
                     factor S-II, Arabidopsis thaliana"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB87703.1"
```

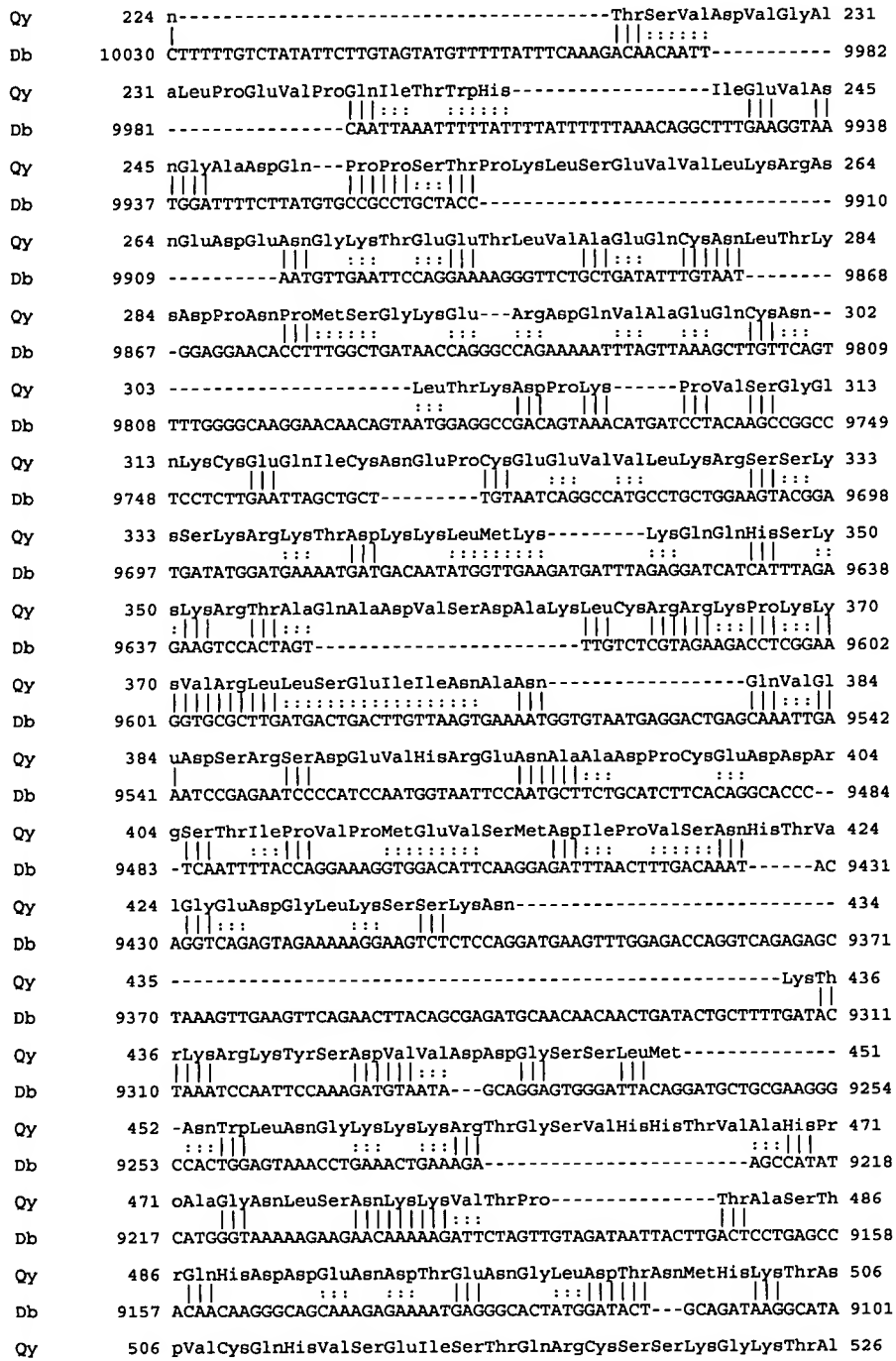
```
                     /db_xref="GI:7573400"
                     /db_xref="GOA:Q9LYE6"
                     /db_xref="InterPro:IPR003618"
                     /db_xref="UniProt/TrEMBL:Q9LYE6"
                     /translation="MAEINGSMQLVGKHKSLPQTTLGGGSASEAPNKQVRPWLQQLSP
                     ASNGILHIPTKILSQETLHSLMHGKKATQTESAPQKPAKPVVNKKQHVPPPQRSVKAM
                     EEVNESVRSKMRESLASALALVKKDDDSPKGKENIGTVETPVITQENTQSFQPASPAS
                     ISVPVGEGTMSEMPTSVESSVQKDSEIPVDIMMEDVIKFNVLKSQYDEVFPRDNVPFT
                     DIIFPNDDLLHGNELSWDLEVSDLGETKDYGTGGEKSFQDPKLLASKIEMELYKLFGG
                     VNKKYRERGRSLLFNLKDKNNPELRERVMSEEISAERLCSMTAEELASKELSQWRQAK
                     AEEMAKMVVLQDTDIDVRSLVRKTHKGEFQVEIEPVDRGTVDVSGGIMSRSKRRPRAK
                     SHSVKTALKDEAAKADNEKSRSTPPSTEEIDPMQGLGIDDELKDVEFLPPIVSLDEFM
                     ESLDSEPPFESPHGNSEMQVSPSEKSDSEAGSDSKSPKGSPKELSDKSLPEAKPEKID
                     EVTPEFDANVKVDDDISRVEKAAALSDDKGERAWDGILQLSMSSVVPVAGIFKSGEKA
                     ETSEWPAMVEVKGRVRLSGFGKFIQELPKSRTRALMVMYLAYKDGISESQRGSLIEVI
                     DSYVADQRVGYAEPASGVELYLCPTRGETLDLLNKVISQEQLDEVKSLDIGLVGVVVW
                     RRAVVPKPGSGSKRQHSFSSSIGSKTSVLPVNKKQRVHVTEKPLVVASMRNHHHGYVK
                     HDTAADDDVPPGFGPVASRDEDDLPEFNFNSSVVPVSSPQPLPAQSKSLDQVRKLIHK
                     YGKSASTYDDDDDEDDIPEWQPHVPSHQLPPPPPPPLGFRPEVFRPPQDGWYDNQNGG
                     SGQHYERNQSRNRGF"
exon                 10281. .11965
                     /gene="F15N18_20"
                     /number=1
intron               11966. .12428
                     /gene="F15N18_20"
                     /number=1
exon                 12429. .12549
                     /gene="F15N18_20"
                     /number=2
intron               12550. .13144
                     /gene="F15N18_20"
                     /number=2
exon                 13145. .13204
                     /gene="F15N18_20"
                     /number=3
intron               13205. .13312
                     /gene="F15N18_20"
                     /number=3
exon                 13313. .14062
                     /gene="F15N18_20"
                     /number=4
gene                 14866. .15413
                     /gene="F15N18_30"
gene                 complement(join(14866. .15009,15090. .15413))
                     /gene="F15N18_30"
CDS                  complement(join(14866. .15009,15090. .15413))
                     /gene="F15N18_30"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB87704.1"
                     /db_xref="GI:7573401"
                     /db_xref="UniProt/TrEMBL:Q9LYE5"
                     /translation="MKPGAFALNPHAASYVPISKRVDYGGGDDGLVFAAKSPTVEVSM
                     PKKSSEMAYKQIRDDDLDLEMDIDMDIEYLLVTFSGLSQESITDVYLANGGDLEATIE
                     MLNQLEIYSTESEENLPETLDIGDISESGPSTSKSTEVAASTSSVIPNAPVSA"
exon                 complement(14866. .15009)
                     /gene="F15N18_30"
                     /number=1
intron               complement(15010. .15089)
                     /gene="F15N18_30"
                     /number=1
exon                 complement(15090. .15413)
                     /gene="F15N18_30"
                     /number=2
gene                 16227. .18163
                     /gene="F15N18_40"
CDS                  join(16227. .16330,16428. .16597,16795. .16924,17004. .17053,
                     17133. .17200,17293. .17388,17469. .17528,17620. .17679,
                     17774. .17828,17930. .17979,18047. .18163)
                     /gene="F15N18_40"
                     /note="the sequence homology is tenuous but is observed
                     over the same sequence stretch for various proteins
                     involved in the water-oxidizing complex of photosystem II
                     similarity to 23 kDa polypeptide of water-oxidizing
                     complex of photosystem II, Nicotiana tabacum,
                     EMBL:NT23WOP2B"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB87705.1"
                     /db_xref="GI:7573402"
                     /db_xref="GOA:P82715"
                     /db_xref="InterPro:IPR002683"
                     /db_xref="UniProt/TrEMBL:P82715"
                     /translation="MRAKGKKQSEEAPEHAIAMALLCPSLPSPNSRLFRSSNISSKYH
                     GASKELMIARSGVSTRSISSEKGLSRRDLVLIGLSSPLSMFLPLSSPVTHAEEDVKMS
                     GEELKMGTMVDDINAYSYAYPLDYPSEKLVFKWVESRKPERYSSAAPLSPDARLRIVS
                     ERVDLTDNLVISISIGPPNSRLTSKEKKTWSAKEVADSVLSDKSALRVTSSQRLEESS
                     VLDAHASDIDGEPYWYYEYLVRKSPTKIAEASKLYRHYISSTAERDGYLYTINASTLG
                     KQWDKGLYKMQMGPVLERAVGSFRLLPPTDSYVPPYKDPWRFW"
exon                 16227. .16330
                     /gene="F15N18_40"
                     /number=1
intron               16331. .16427
                     /gene="F15N18_40"
                     /number=1
exon                 16428. .16597
                     /gene="F15N18_40"
                     /number=2
intron               16598. .16794
                     /gene="F15N18_40"
                     /number=2
exon                 16795. .16924
                     /gene="F15N18_40"
                     /number=3
intron               16925. .17003
                     /gene="F15N18_40"
                     /number=3
exon                 17004. .17053
                     /gene="F15N18_40"
                     /number=4
intron               17054. .17132
                     /gene="F15N18_40"
                     /number=4
exon                 17133. .17200
                     /gene="F15N18_40"
                     /number=5
intron               17201. .17292
                     /gene="F15N18_40"
                     /number=5
exon                 17293. .17388
                     /gene="F15N18_40"

Alignment Scores:
Pred. No.:              4.81e-05      Length:          101665
Score:                  298.00        Matches:         225
Percent Similarity:     35.5%         Conservative:    152
Best Local Similarity:  21.2%         Mismatches:      360
Query Match:            5.4%          Indels:          324
DB:                     15            Gaps:            50

US-09-828-068-2 (1-1057) x ATF15N18 (1-101665)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
              :::|||     |||     :::   |||     |||   |||       |||
Db      59795 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 59845

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
                |||   |||::::::           |||:::|||:::   :::       |||
Db      59846 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 59905

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla-------------------- 247
                  :::       :::       |||:::
Db      59906 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 59965

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
                |||   ||||||         |||     :::   :::           |||
```

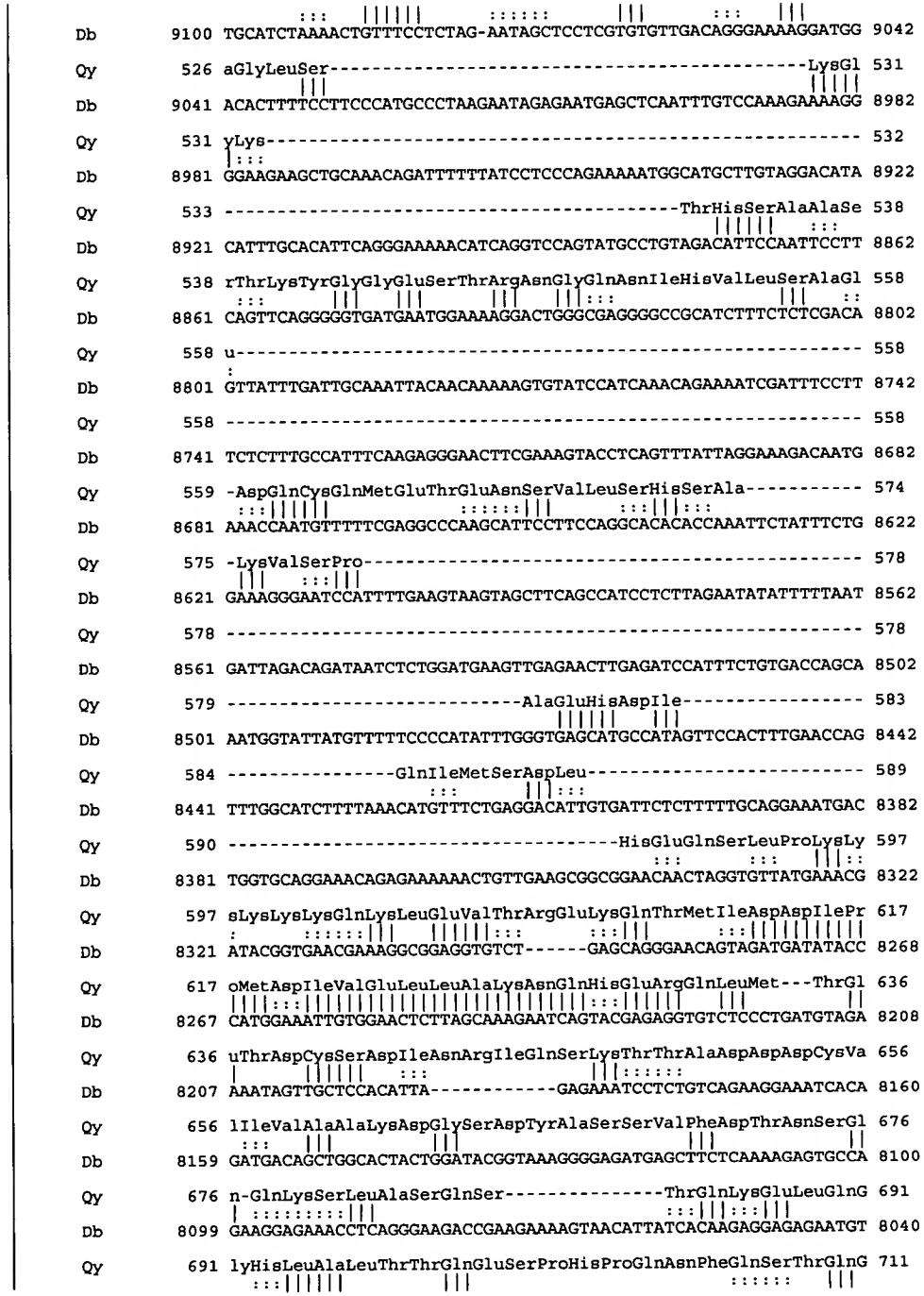

Db 59966 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 60025
Qy 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db 60026 GATAACGGGTCATCAGAAAGTATAAATCTTGCT---------- 60058
Qy 287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
Db 60059 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 60112
Qy 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db 60113 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA---------- 60154
Qy 324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db 60155 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 60208
Qy 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
Db 60209 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 60268
Qy 361 AlaLysLeuCys---------- 364
Db 60269 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 60328
Qy 365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db 60329 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 60388
Qy 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db 60389 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 60442
Qy 400 CysGluAspAspArgSerThr----------IlePro 408
Db 60443 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 60496
Qy 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db 60497 TGTCCT----------CCGGGTACTCAGAGAACA------GAGAGGAAG 60529
Qy 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db 60530 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 60580
Qy 449 SerLeuMet----------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
Db 60581 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 60640
Qy 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db 60641 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 60700
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
Db 60701 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 60760
Qy 500 ----------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db 60761 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 60820
Qy 516 ThrGlnArgCys----------SerSerLysGlyLysThrAlaGly---Leu 528
Db 60821 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 60880
Qy 529 SerLysGly----------LysThr 533
Db 60881 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 60940
Qy 534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db 60941 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 61000
Qy 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db 61001 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 61060
Qy 573 r---------- 573
Db 61061 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 61120
Qy 574 ----------Ala-LysValSerProAlaGluHisAspIleG 584
Db 61121 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 61180
Qy 584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Db 61181 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 61240
Qy 604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db 61241 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 61297
Qy 624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db 61298 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 61357
Qy 643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db 61358 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 61417
Qy 663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
Db 61418 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 61477
Qy 681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db 61478 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 61528
Qy 701 roHis----------ProGlnAsn---------PheGlnS 708
Db 61529 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 61588
Qy 708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db 61589 CTACCCAAGAAAACCGA----------GCCAGCT 61612
Qy 728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db 61613 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 61666
Qy 748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db 61667 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT---------- 61710
Qy 768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
Db 61711 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------- 61746
Qy 785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Db 61747 --GAAGCTTCTCATCCAATTTGGCCATCTTCC----------ATGATAC 61783
Qy 805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db 61784 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 61840
Qy 825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db 61841 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 61894
Qy 845 alLeuTyrProLysGlu---------- 850
Db 61895 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 61954

```
Qy        851 ----------------------------------------------------------SerM 852
                                                                        |||:
Db      61955 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 62014

Qy        852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
                ::||||||  |||||||| ::::::||||||          ::::::   |||
Db      62015 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 62074

Qy        871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
                   :::|||  :::    ::::::                       :::   :
Db      62075 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 62134

Qy        888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
              :::::         |||          |||:::  ::::::|||   :::
Db      62135 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 62187

Qy        908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
                 :::||||||      |||:::      |||:::             :::     |||
Db      62188 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 62239

Qy        928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
                      |||       |||    :::|||          |||   |||     :::   |
Db      62240 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 62293

Qy        947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
              ||:::||||||   :::
Db      62294 CTTCATGGAGTCCTCAT------------------------------------------- 62310

Qy        967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
                        ||||||     :::   :::      :::                    |||:
Db      62311 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC---------------TTTG 62347

Qy        987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
              ::       :::|||        :::                :::||||||   |||
Db      62348 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 62407

Qy       1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
                        |||    |||    ||||||                                |
Db      62408 TCCAGCTGCTGGGAGCATCGAATTCC----------------------------------A 62434

Qy       1022 etValHisProLeu---------------AspArgPheValArgGlnAsp---------- 1033
              ||:::    ||||||                |||:::       :::|||
Db      62435 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 62494

Qy       1034 ----------------------------------------------------------IleC 1035
                                                                        :::|
Db      62495 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 62554

Qy       1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
               ||        |||:::|||||||||||||||||              |||    ||||||
Db      62555 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 62613
```

RESULT 12
AP006698/c

```
LOCUS       AP006698               108174 bp    DNA     linear   PLN 03-FEB-2004
DEFINITION  Lotus corniculatus var. japonicus genomic DNA, chromosome 1,
            clone:LjT04O06, TM0487, complete sequence.
ACCESSION   AP006698
VERSION     AP006698.1  GI:41688383
KEYWORDS    HTG.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Kato,T., Sato,S., Nakamura,Y., Kaneko,T., Asamizu,E. and Tabata,S.
  TITLE     Structural Analysis of a Lotus japonicus Genome. V. Sequence
            Features and Mapping of Sixty-four TAC Clones Which Cover the 6.4
            Mb Regions of the Genome
  JOURNAL   DNA Res. 10, 277-285 (2003)
REFERENCE   2  (bases 1 to 108174)
  AUTHORS   Sato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-2003) Shusei Sato, Kazusa DNA Research Institute,
            Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
            Chiba 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
            Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1..108174
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /chromosome="1"
                     /clone="LjT04O06"
                     /clone_lib="LjT library"
                     /note="TAC clone:TM0487, synonym:Lotus japonicus"
ORIGIN

Alignment Scores:
Pred. No.:              0.00186        Length:         108174
Score:                  269.00         Matches:        258
Percent Similarity:     31.2%          Conservative:   167
Best Local Similarity:  18.9%          Mismatches:     463
Query Match:            4.9%           Indels:         478
DB:                     15             Gaps:           56

US-09-828-068-2 (1-1057) x AP006698 (1-108174)

Qy         63 GlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIlePhe 82
              ||||||||||||      ::::::|||||||||      |||   |||         :::
Db      10675 GGATATGTGGCTGAAATGCGCAAGAAAGACTGGAAACTTTGTTGGCCATTCCCAGTACTT 10616

Qy         83 HisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerValAlaLys 102
                       |||                           |||      |||    |||
Db      10615 GACTCTGACAAG---------------CCACCAATCCTTCCACCATTAGAGGTACGGAAA 10571

Qy        103 PheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThr 122
                |||:::   |||    |||          |||   ::::::      :::|||
Db      10570 TTCAAATGTTGGCGTTGCCCAGATGTTCTACAAGAAGTGGCTGGCGCAGACTTTCATAAA 10511

Qy        123 AlaProArgThr------------------------------------------------ 126
                    :::|||
Db      10510 GATGATCAAACAGACTTAAATTGTCATAGTACTGAGTGCAAATCTGATCCCAACTGTAGT 10451

Qy        127 ---LeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg--- 144
                    |||   |||                   |||             |||   |||
Db      10450 AATGCGCCTCTCAAGTCTGGTAATCAGCAAGATCCAGTGCCTGACACTCATGAGAGAGAA 10391

Qy        145 ------------SerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
                          :::|||   :::              |||          :::|||
Db      10390 GAAATTGATCTTAATACAACTCTTAGCTGTCTCGATGGTTGCTTACCAATTAGTAATGAG 10331

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArg--------------------------- 171
                 :::          |||         :::|||
Db      10330 AAAGAAAAGGAAGCCGGTGTTACACTTAACAGAATAACAGGTAACTATCAGCATCAGTTA 10271

Qy        172 ---------------------------------------------SerThrLeuProLys 176
                                                            |||      |||
Db      10270 TATAAGAGTATCTTGTTTATGTTTTTACAGTAGGTGGTTTGATTTTCTGTATTTCCTCCA 10211

Qy        177 SerValGlnGluGly---------------AsnAspSerLysCysAsnAlaProSerGly 191
                ::::::    |||                          |||         |||   ||||||
Db      10210 GAGATTGAGGTTGGCTTGGGTGATAATCTCAATCACCAAGATTCTAATGTCCCATCATCT 10151

Qy        192 LysAsn------GlyAlaAlaGluAlaAsn-ThrAspSerProMetLysAspLeuGlnGl 209
              |||:::         |||   |||:::           |||             |||
Db      10150 AAAAGTTATCCAGGCCTTGCCCAAGAAATGCACACAACTAACAATGGTAAGTTCAAACTG 10091

Qy        209 yProAlaGlnAsn---------------TyrAspValAlaAlaAsnValSerGluAspAs 224
                   |||   |||                    |||     :::                |||||
Db      10090 GACAGCACATAATTCAGTCAAGCAGCTGTACTTCATCCTTTTTACTATTGAAGTGACAA 10031
```





```
Db       8039 GAAACCTAGCAAAAGAAATTCAGCTCATCACTTTCCTTCATTTGATGGAAACAATTCAGG 7980

Qy        711 luGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuP 731

Db       7979 TGTGAATAATCTATATCCACCTCAATCACCCTATGGATTTGAAGTTTCTCAGTCCCAAAA 7920

Qy        731 heSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysA 751

Db       7919 GAGGCCATCAAGTGGATTCCAA-TTTTCACCGGTTTCTAGCCAGCTTGGCAGTGCTAAAA 7861

Qy        751 spAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAlaT 771

Db       7860 AGTGCAAGTTTAGTGGGAGTCTTGAAGAGCGTGCACCCTCTAATGCTGCTTGGCAAGGCC 7801

Qy        771 hrCysGlyAlaGlnPheArgProGlyIle---------------GlnAlaValAspLeuT 786

Db       7800 TGGGAGGATGTAGCTTACACAAGAATATCTTATTACAGCAGAATGAATGGGCATCCTTGA 7741

Qy        786 hrSerThrHisVal---MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805

Db       7740 CCTCAAATCATGTTTCCATAGGGCATGTCCCCCAAAAGGTTGTTTCTCAGCCTACTAATA 7681

Qy        805 laProLeuAspArgTyrAlaGluArgAla------------------------------ 814

Db       7680 GTAACATGGACATAGCTTCACTCCGATCTGGTCTGCTTCATAAGCAGAACATGAAGAGGG 7621

Qy        815 ----------------------ValAsn------------GlnValHisAlaArgAsn- 822

Db       7620 ATATTGATCTGAACTACATAAATATCAATGCTGCTGGCTTAGAAAAGCATAGTAGGAATG 7561

Qy        823 ----------------------------------------PheProSerThrIleAlaT 829

Db       7560 CAGGTACTGGGACCTTCAGCAGGATGAATGGAGAATACACATTTCCTTGTAAACATAATG 7501

Qy        829 hrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProL 849

Db       7500 GAATGGAGCCT------------CATCAAAATTCGAGGGGATCCCTGGATTTTTATTCTA 7453

Qy        849 ysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerP 869

Db       7452 ACGAGACAATACCTGCCATGCATTTACTCAGCCTTATGGATGCAGGCATGCAGACGCGGA 7393

Qy        869 hePro---AsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888

Db       7392 CGCCCTTCAATGTGGGAGCAAGTGCT------CAAATGCTTAAAAGA------------- 7352

Qy        888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908

Db       7351 ----------------------------------CCCTCTTATCCCGGTGATTGCAATA 7327

Qy        908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAsp---LeuHisA 927

Db       7326 CTAAGTTG--------------GAGATTGGTACATCTAAGGCCCATGATACCCTGAAAA 7282

Qy        927 rgPro--------------------------LeuArgProHisProArgValGlyValL 938

Db       7281 GGCCACCACCTGATTATTACAGCAGAAGTTACTTATCTGATAAACCCCATGGTTGTTTTA 7222

Qy        938 euGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerG 958

Db       7221 TTGGCTCT-----------------CCAACCTTTGGTGCTTCTGCATCAACCCAGCCCG 7180

Qy        958 lyTyrLysLeuGlyValSerThrGlyIle---------ThrSerHisGlnMetAsnArgL 975

Db       7179 ATAAGAACTTAATAGGGGCTACCGGTTTCAATGGTGAAAATTCAACAAAGTTGGGAAAGA 7120

Qy        975 ysGluHisPheGluAlaLeuAsnSer---------GlyMetPheSerAlaLysTrpAsnA 992

Db       7119 AAGAAAAGATGAAGAGCTCCAACTCAACCCTGCAGAATGTAGTTAGTAAGCCGTTTAGCT 7060

Qy        992 laLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaArgAsnSerIleA 1012

Db       7059 TGCCTCGTCTTGAAATGGAAACTTCACTGCAGCACAAATTAGACGCT------------- 7013

Qy       1012 laGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPheValArgG 1032

Db       7012 -------------CATGGCACTCAAGAAACCCCAGTGCCGTTAAAT---GTCATCTCTG 6970

Qy       1032 lnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnG 1052

Db       6969 GAAACTCATGTATGGTGAACAGAAATCCAGCTGACTTCACCACACCCGAAACAGGAAATG 6910

Qy       1052 luTyrMet 1054

Db       6909 TCTATATG 6902
```

RESULT 13
AY605097

```
LOCUS       AY605097               5397 bp    DNA     linear   SYN 12-JUL-2004
DEFINITION  Synthetic construct fusion of carotene synthesis proteins
            (crtEBIWY) gene, complete cds.
ACCESSION   AY605097
VERSION     AY605097.1  GI:47531117
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 5397)
  AUTHORS   Xiong,A.S., Yao,Q.H., Peng,R.H., Li,X., Fan,H.Q., Cheng,Z.M. and
            Li,Y.
  TITLE     A simple, rapid, high-fidelity and cost-effective PCR-based
            two-step DNA synthesis method for long gene sequences
  JOURNAL   Nucleic Acids Res. 32 (12), E98 (2004)
   PUBMED   15240836
REFERENCE   2  (bases 1 to 5397)
  AUTHORS   Yao,Q.-H., Xiong,A.-S. and Peng,R.-H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2004) Agro-Biotechnology Research Center,
            Shanghai Academy of Agricultural Sciences, Beidi Road 2901,
            Shanghai 201106, China
FEATURES             Location/Qualifiers
     source          1. .5397
                     /organism="synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /note="derived from Paracoccus marcusii"
     gene            1. .5397
                     /gene="crtEBIWY"
     CDS             1. .5397
                     /gene="crtEBIWY"
                     /note="CrtE/CrtB/CrtI/CrtW/CrtZ"
                     /codon_start=1
                     /transl_table=11
                     /product="fusion of carotene synthesis proteins"
                     /protein_id="AAT35222.1"
                     /db_xref="GI:47531118"
                     /translation="MSDLVLTSTEAITQGSQSFATAAKLMPPGIRDDTVMLYAWCRHA
                     DDVIDGQALGSRPEAVNDPQARLDGLRADTLAALQGDGPVTPPFAALRAVARRHDFPQ
                     AWPMDLIEGFAIDVEARDYRTLDDVLEYSYHVAGILGVMMARVMGVRDHPVLDRACDL
                     GLAFQLTHIARDVIDYARIGRCYLPGDWLDHAGARVDGPVPSPELYTVILRLLDAAEP
                     YYPSARVGLADLPPRCAWSIPAAFRIYRAIGLPIRKGGPEAYRQRISTSKAAKIGLLG
                     IGGWDVADHACRGSGVSRQDLWTRPHHAQLLNFDLLKLAGDVESNPGPMRRDVNPIHA
                     TLLQTRIEEIAQGFGAVLQPLGAAMAALSSGKRFRGMLMLLAAEASGGVCDTIVDAAC
                     AVEMVHAASLIFDDLPCMDDAGLRRGQPATHVAHGESRAVLGGIALITEAMALLAGAR
                     GASGTVRAQLVRILSRSLGPQGLCAGQDLDLHAAKNGAGVEQEQDLKTGVLFIAGLEM
                     LAVIKEFDAEEQTQMIDFGRQLGRVFQSYDDLLEVVGDQAALGKDTGRDAAAPGPRRG
                     LLAVSDLQKVSRQYEASRAQLEGMVGSKRLQAPEIEALLERVLPYAARAQLLNFDLLK
                     LAGDVESNPGPMNAHSPAAKTAIVIGAGFGGLALAIRLQSAGIATTLVEARDKPGGRA
                     YVWHDQGHVFDAGPTVITDPDALKELWALTGQDMARDVTLMPVSPFYRLMWPGGKVFD
                     YVNEADQLERQIAQFNPDDLEGYRRFRDYAEEVYQEGYVKLGTVPFLKLGQMLKAAPA
                     LMKLEAYNSVHAKVATFIKDPYLRQAFSYHTLLVGGNPFSTSSIYALIHALERRGGVW
                     FAKGGTNQLVAGMVALFERPGGQMMLNAKVARIETEGARTTGVTLADGRSLRADMVAS
                     NGDVMHNYRDLLGHTARGQSRAKSLDRKRWSMSLFVLHFGLREAPKDIAHHTILFGPR
                     YKELVNEIFKGPKLAEDFSLYLHSPCTTDPDMAPPGMSTHYVLAPVPHLSRAEIDWAV
                     EGPRYADRILAFLEERLIPNLRANLTRTRIFTPADFASELNAHHGSAFSVEPILTQSA
                     WFRPHNRDKTIRNFYLVGAGTHPGAGIPGVVGSAKATAQVMLQLLNFDLLKLAGDVES
```

```
                     NPGPMSAHALPKADLTATSLIVSGGIIAAWLALHVHALWFLDAAAHPILAVANFLGLT
                     WLSVGLFIIAHDAMHGSVVPGRPRANAAMGQLVLWLYAGFSWRKMIVKHMAHHRHAGT
                     DDDPDFDHGGPVRWYARFIGTYFGWREGLLLPVIVTVYALILGDRWMYVVFWPLPSIL
                     ASIQLFVFGTWLPHRPGHDAFPDRHNARSSRISDPVSLLTCFHFGGYHHEHHLHPTVP
                     WWRLPSTRTKGDTAQLLNFDLLKLAGDVESNPGPMTHDVLLAGAGLANGLIALALRAA
                     RPDLRVLLLDHAAGPSEGHTWSCHDPDLSPHWLARLKPLRRANWPDQEVRFPRHARRL
                     ATGYGSLDGAALADAVVRSGAEIRWDSDIALLDAQGATLSCGTRIEAGSVLDGRGPHP
                     SRHLTLGFHKFLGVEIETDRPHGVPRPVIMDGTVTQRDGYGFIYLLPFSRTRILIEDT
                     RYSDGGDLDDDALAAASQDYARQQGWTGAEVRRERGILPIALAHDAAGFWADHAEGPV
                     PVGLRAGFFHPVTGYSLPYAAQVADVVAGLSGPPGTDALRGAIRDYAIDRARRDRFLR
                     LLNRMLFRGCAPDRRYTLLQRLYRMPHGLIERFYAGRLSVADQLRIVTGKPPIPLGTA
                     IRCLPERPLLKENA"
ORIGIN

Alignment Scores:
Pred. No.:              0.000506        Length:         5397
Score:                  251.00          Matches:        199
Percent Similarity:     34.3%           Conservative:   145
Best Local Similarity:  19.9%           Mismatches:     375
Query Match:            4.5%            Indels:         284
DB:                     11              Gaps:           42

US-09-828-068-2 (1-1057) x AY605097 (1-5397)

Qy       121 GlyThrAlaProArgThrLeu---------ProAlaLysGlnAsnGlyThrSerAspGly 137
Db       839 GGGACGTCGCAGACCACGCTTGCAGAGGCTCCGGTGTCTCCAGACAAGACCTCTGGACCA 898

Qy       138 CysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysVal 157
Db       899 GACCACACCACGCACAACTCCTCAACT------TCGACCTCCTCAAACTCGCAGGCGACG 952

Qy       158 SerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSer 177
Db       953 TCGAATCCAACCCAGGCCCAATGAGAAGAGACGTCAACCCAATCCACGCAACCCTCCTCC 1012

Qy       178 ValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGlu 197
Db      1013 AAACCAGAATCGAAGAAATCGCACAAGGCTTCGGTGCAGTCCTCCAACCACTCGGTGCAG 1072

Qy       198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db      1073 CAATGGCAGCACTCTCCT-----------------CCGGTAAGAGATTCAGAGGTATGC 1114

Qy       218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
Db      1115 TCATGCTCCTCGCAGCAGAAGCATCCGGTG---GTGTCTGCGACACCATCGTCGATGCAG 1171

Qy       238 ----------------------------------------IleThrTrpHisIle--- 242
Db      1172 CCTGCGCAGTCGAAATGGTCCACGCAGCATCCCTCATCTTCGACGACCTCCCATGTATGG 1231

Qy       243 -----------------GluValAsnGlyAlaAspGlnProProSer---------Thr 253
Db      1232 ACGACGCAGGTCTCAGAAGAGGTCAAC------CAGCAACCCACGTCGCACACGGTGAAT 1285

Qy       254 ProLysLeuSerGluValValLeuLysArgAsnGluAspGlu------AsnGlyLysThr 271
Db      1286 CCAGAGCAGTCCTCGGTGGCATCGCACTCATCACCGAAGCAATGGCACTCCTCGCAGGTG 1345

Qy       272 GluGluThr------LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMet 289
Db      1346 CAAGAGGTGCATCCGGTACTGTCAGAGCACAACTCGTCAGAATCCTCTCCAGATCCCTCG 1405

Qy       290 SerGlyLysGluArgAspGlnVal----------------------------AlaGlu 299
Db      1406 GTCCACAAGGTCTCTGCGCAGGTCAAGACCTCGACCTCCACGCAGCAAAGAACGGTGCAG 1465

Qy       300 GlnCysAsnLeuThrLysAspProLysProValSerGly------------GlnLysCys 315
Db      1466 GTGTCGAACAAGAACAAGACCTCAAAACCGGTGTCCTCTTCATCGCAGGTCTCGAAATGC 1525

Qy       316 GluGlnIleCysAsnGluProCysGluGlu---------------------------- 325

Db      1526 TCGCAGTCATCAAAGAGTTCGACGCAGAAGAACAAACCCAAATGATCGACTTCGGTAGAC 1585

Qy       326 --------------------------ValValLeuLysArgSerSerLysSerLysArg 336
Db      1586 AACTCGGTAGAGTCTTCCAATCCTACGACGACCTCCTCGAAGTCGTCGGTGATCAAGCAG 1645

Qy       337 LysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAla 356
Db      1646 CACTCGGTAAAGACACCGGTAGAGACGCAGCAGCACCAGGTCCAAGAAGAGGTCTCCTCG 1705

Qy       357 AspValSerAspAlaLysLeuCysArgArgLysProLys---------------------- 369
Db      1706 CA-GTCTCCGACCTCCAAAAGGTCTCCAGACAATACGAAGCATCCAGAGCACAACTCGAA 1764

Qy       370 ------------LysValArgLeuLeuSerGluIleIleAsnAla-AsnGlnValGluAs 385
Db      1765 GGTATGGTCGGTTCCAAGAGACTCCAAGCACCAGAAATCGAAGCACTCCTCGAAAGAGTC 1824

Qy       385 pSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSe 405
Db      1825 CTCCCATACGCAGCAAGAGCACAACTCCTCAACTTCGACCTCCTCAAACTCGCAGGCGAC 1884

Qy       405 rThrIleProVal-----------------------------------ProMetGluVa 413
Db      1885 GTCGAATCCAACCCAGGCCCAATGAACGCACACTCCCCAGCAGCAAAGACCGCAATCGTC 1944

Qy       413 lSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLy 433
Db      1945 ATCGGTGCAGGTTTCGGTGGTCTCGCACTCGCAATCAGACTCCAATCCGCAGGCATCGCA 2004

Qy       433 sAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetAsnTr 453
Db      2005 ACCACCCTCGTCGAGGCAA-----------------GAGACAAGCCAGGCGGTAGAGCA 2046

Qy       453 pLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisProAlaGl 473
Db      2047 TACGTCTGGCACGACCAAGGCCACG---------------------------------- 2071

Qy       473 yAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGluAsnAs 493
Db      2072 -------TCTTCGACGCAGGTCCAACCGTCATCACCGACCCAGACGCACTCAAGGAACTC 2124

Qy       493 pThrGluAsnGlyLeuAspThr--------------------------AsnMetHisLy 504
Db      2125 TGGGCACTCACCGGTCAAGACATGGCAAGAGACGTCACCCTCATGCCAGTCTCCCCATTC 2184

Qy       504 sThrAspValCysGlnHisVal---------SerGluIleSerThrGlnArgCysSerSe 521
Db      2185 TACAGACTCATGTGGCCAGGTGGCAAGGTCTTCGACTACGTCAACGAGGCAGACCAACTC 2244

Qy       521 rLysGlyLysThr------------------------------------------------ 525
Db      2245 GAAAGACAAATCGCACAATTCAACCCAGACGACCTCGAAGGTTACAGAAGATTCAGAGAC 2304

Qy       526 ------------------------AlaGlyLeuSerLysGlyLysThrHisSerAlaAl 537
Db      2305 TACGCAGAGGAGGTCTACCAGGAGGGCTACGTCAAACTCGGCACCGTCCCATTCCTCAAG 2364

Qy       537 aSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAl 557
Db      2365 CTCGGTCAAATGCTCAAGGCAGCACCAGCACTCATGAAACTCGAAGCATACAACTCCGTC 2424

Qy       557 aGluAspGlnCysGln---------------------------------------------- 562
Db      2425 CACGCAAAGGTCGCAACCTTCATCAAGGACCCATACCTCAGACAGGCATTCTCCTACCAC 2484

Qy       563 ------------------------------------------------------MetGl 564
Db      2485 ACCCTCCTCGTCGGTGGTAATCCATTCTCCACCTCCTCCATCTACGCACTCATCCACGCA 2544

Qy       564 uThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGluHisAspIleGl 584
```

```
Db       2545 CTCGAAAGAAGAGGTGGTGTCTGGTTCGCAAAGGGTGGCACCAACCAGCTCGTCG---CA 2601

Qy        584 nIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGl 604

Db       2602 GGTATGGTCGCACTCTTCGAGAGACCAGGTGGTCAAATGATGCTCAACGCAAAGG----- 2656

Qy        604 uValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLe 624

Db       2657 ----TCGCAAGAATCGAAACCGAGGGTGCAAGAACCACCGGTGTCACCCTCGCAGACGGT 2712

Qy        624 uAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer---------- 640

Db       2713 AGATCCCTCAGAGCAGACATGGTCGCATCCAACGGTGACGTCATGCACAACTACAGAGAC 2772

Qy        641 ------------------AspIleAsnArgIleGlnSerLysThrThr---AlaAspAs 653

Db       2773 CTCCTCGGTCACACCGCAAGAGGTCAATCCAGAGCAAAGTCCCTCGACAGAAAGAGATGG 2832

Qy        653 pAspCysValIleValAlaAlaLysAspGlySerAspTyr---------- 666

Db       2833 TCCATGTCCCTCTTCGTCCTCCACTTCGGTCTCAGAGAAGCACCAAAGGACATCGCACAC 2892

Qy        667 -------AlaSerSerValPheAspThr---AsnSerGlnGlnLysSerLeuAlaSerGl 683

Db       2893 CACACCATCCTCTTCGGTCCAAGATACAAGGAACTCGTCAACGAAATCTTCAAGGGTCCA 2952

Qy        683 nSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisPr 703

Db       2953 AAACTCGCAGAGGACTTCTCCC---------TCTACCTCCACTCCCCATGCACCACCGAC 3003

Qy        703 oGlnAsnPheGln----------SerThrGlnGluGlnGlnThrHisLeuAr 717

Db       3004 CCAGACATGGCACCACCAGGCATGTCCACCCACTACGTCCTCGCACCAGTCCCACACCTC 3063

Qy        717 gMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGl 737

Db       3064 TCCAGAGCAGAAATCGACTGGGCAGTCGAAGGCC---------------CAAGATACGCA 3108

Qy        737 nTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTr 757

Db       3109 G------------ACAGAATCCTCGCATTCCTCGAAGAAAGACTCATCCCAA---ACCTC 3153

Qy        757 pGluGlnPheLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheAr 777

Db       3154 AGAGCAAACCTCACCAGAACCAGAATCTTCACCCCAGCAGACTTCG------------CA 3201

Qy        777 gProGlyIleGlnAlaValAspLeuThrSerThrHisValMetGly------SerSerSe 795

Db       3202 TCCG--------------------AACTCAACGCACACCACGGCTCCGCATTCTCCGTC 3240

Qy        795 rAsnTyrAlaSerArgGlnPro---------------------------------- 802

Db       3241 GAACCAATCCTCACCCAATCCGCATGGTTCAGACCACACAACAGAGACAAGACCATCAGA 3300

Qy        803 ----ValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaAr 821

Db       3301 AACTTCTACCTCGTCGGTGCAGGTACTCACCCAGGTGCAGGTATCCCAGGTG-------- 3352

Qy        821 gAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAl 841

Db       3353 ----------TCGTCGGTTCCGCAAAGGCAACCGCACAAGTCATGCTCCAACTCCTCAAC 3402

Qy        841 a--------------GlyGlnValValLeuTyrProLysGluSerMet---ProAlaTh 855

Db       3403 TTCGACCTCCTCAAACTCGCAGGCGACGTCGAATCCAACCCAGGCCCAATGTCCGCACAC 3462

Qy        855 rHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro------------- 870

Db       3463 GCACTCCCAAAAGCAGACCTCACCGCAACCTCCCTCATCGTCTCCGGTGGTATCATCGCA 3522

Qy        871 -------AsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888

Db       3523 GCATGGCTCGCACTCCACGTCCACGCACTCTGGTTCCTCGACGCAGCAGCAC---ACCCA 3579
```

```
Qy        888 nTyrAlaHisAsnGlnTyrLys-----------GlySerThrSerThrSerTyrGlySe 904

Db       3580 ATCCTCGCAGTCGCAAACTTCCTCGGTCTCACCTGGCTCTCCGTCGGTCTCTTCATCATC 3639

Qy        904 rAsn 905

Db       3640 GCAC 3643

RESULT 14
CR848207
LOCUS       CR848207               3144 bp    mRNA    linear   VRT 12-OCT-2004
DEFINITION  Xenopus tropicalis finished cDNA, clone TGas105h10.
ACCESSION   CR848207
VERSION     CR848207.1  GI:54110308
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 3144)
  AUTHORS   Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
            Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
            McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
            Voigt,J. and Zorn,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT     Sanger Xenopus tropicalis EST/cDNA project.
            This sequence is from a Xenopus Gene Collection (XGC) library, from
            a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
            extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
            ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .3144
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas105h10"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stage 10.5-13 mixed)"
ORIGIN

Alignment Scores:
Pred. No.:                0.000389        Length:          3144
Score:                    248.00          Matches:         214
Percent Similarity:       34.6%           Conservative:    160
Best Local Similarity:    19.8%           Mismatches:      425
Query Match:              4.5%            Indels:          283
DB:                       5               Gaps:            42

US-09-828-068-2 (1-1057) x CR848207 (1-3144)
```

```
Qy         71 LysAspProLysPheCys------SerLeuSerArgIlePheHisAspGlnLysLysCys 88

Db        136 AAAGACAATGATTCCTGTCCAACATCTCTTCTGAATATATTTAGCGACTGGATGAAGTCA 195

Qy         89 AspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCys 108

Db        196 CCAGATGCAAAG---------------------------------------------- 207

Qy        109 SerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128

Db        208 --------------AAGAGAAAGCCATCAGCTAATGGT---TTGCCTAAGAAAAAGCCT 249

Qy        129 AlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheVal 148

Db        250 ACAAACGAGTCTTCCAGCAGCGAAGACAGCTCC---------------------------- 282
```

```
Db       2194 ACA--------------AAAACAAATACTGCTATTTCTAAATCCCCAGTTACTCCC--- 2235

Qy        808 AspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIle 827

Db       2236 --------------------------------AAAGCAAAACCTGCTCCTAAAAAAGAA 2262

Qy        828 AlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGly-GlnValValLeuTy 847

Db       2263 AGCAGCAGTGAGAGTTCTGATTCTGAAGATGAAAAGCAAAAAGGAAAGAATATCTCTACT 2322

Qy        847 rProLysGluSerMetPro------------AlaThrHisLeuLeuArgMetMetAspPr 863

Db       2323 CCCAAAACAGCAAATTCCACTCCTAAAGCAGCTGCTGAGAGTTCTGAGGAAAGTTCTTCC 2382

Qy        863 oSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGl 883

Db       2383 AGTGAGGATGAAGAAAAA-GCTAAC---GGGACCTCTGGAAAACGGAAAAGAGGGGAAAG 2438

Qy        883 nLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGl 903

Db       2439 TACAGGCAATGCAGAATGT--------------------GAAACTATGACACCAGAAAA 2477

Qy        903 ySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHi 923

Db       2478 CAAGAAATTGAAAGCTAAAACTCCAAACACATTTCCCAAGGTTAACAAGAAGGAA---CA 2534

Qy        923 sAspLeuHisArgProLeuArgPro------------------------HisProArgVa 935

Db       2535 CAACAAAAATACTCCATTTCGGAGGGTGAGAGAAGAAGAAGTAGAAATCAATCCTCGGAT 2594

Qy        935 l--------------------------GlyValLeuGlySerLeuLeuGlnLysGluIl 946

Db       2595 GACTGATAATGCATTTGAAGCAAAGAAAGGAGCACAAGGA------------------- 2634

Qy        946 eAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGly----------- 962

Db       2635 ----GACTGGGGTGAAAAGGCCAATAATGTCCTTAAATTCACAAAGGGAAAATCATTCAG 2690

Qy        963 ---------------------------------------------ValSerThrGlyIl 967

Db       2691 ACACGAGAAGACCAAGAAGAAACGTGGCAGTTACTGCGGAGGTGCAATTTCCACATCTGT 2750

Qy        967 eThrSerHisGlnMetAsnArg------LysGluHisPheGluAlaLeuAsnSerGlyMe 985

Db       2751 CAATTCAATCAAGTTTGACAGTGATTAAAAGGATTGTAAGGAAGAACTTTTTCCTGGTTC 2810

Qy        985 tPheSerAla--------------LysTrpAsnAlaLeuGlnLeu-GlySerValSerS 1000

Db       2811 AACCACAAGCACACGAGCTAGGGAACGGTGGCATCATTTAAGTCTGAGCGAATTCCGACG 2870

Qy       1000 erSerAlaAspPheLeuSerAlaArgAsnSerIle------AlaGlnSerTrpThrArg 1017

Db       2871 AGTCGATACGTTTATTAGCAGTTTCCAGCAGCATTTTTGCAGTTCAACAATGGACGAAA 2929
```

RESULT 15
BC064870

```
LOCUS       BC064870               3154 bp    mRNA    linear   VRT 23-AUG-2004
DEFINITION  Xenopus tropicalis hypothetical protein LOC395027, mRNA (cDNA clone
            IMAGE:5335805), partial cds.
ACCESSION   BC064870
VERSION     BC064870.1  GI:40675334
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 3154)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3154)
  AUTHORS   Klein,S. and Gerhard,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2004) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
  REMARK    NIH-MGC Project
COMMENT     Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Robert M. Grainger
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 142 Row: f Column: 5
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES              Location/Qualifiers
     source           1. .3154
                      /organism="Xenopus tropicalis"
                      /mol_type="mRNA"
                      /db_xref="taxon:8364"
                      /clone="IMAGE:5335805"
                      /tissue_type="Embryo, Silurana tropicalis, tailbud (stages
                      20-27)"
                      /clone_lib="NICHD_XGC_Emb7"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6.1"
     gene             <1. .3154
                      /gene="LOC395027"
                      /db_xref="GeneID:395027"
     CDS              <1. .2752
                      /gene="LOC395027"
                      /codon_start=2
                      /product="LOC395027 protein"
                      /protein_id="AAH64870.1"
                      /db_xref="GI:40675335"
                      /db_xref="GeneID:395027"
                      /translation="DGSSVVPSDLFPHVLRFLRENKFTNAANEFSKATGVKDNDSCPT
                      SLLNIFSDWMKSPDAKKRKPSANGLPKKKPTNESSSSEDSSSEEDEPPAKKIATQSAE
                      GKKPVVKGVQAKKAKSSSEDSSDESDSEKETKKPPVKKPAQTPKGAALKTPTQKKAES
                      SSSESSSSEDEASKKKKQSVTKVPPKQTVVKAGLANDNQKAADSSSSEDSDSPPEKKS
```

AAPKTPPAKPATAAKPQAKKTAGKKSSSEDSSDSSDEEAKPAKSKPKPGVYSAVPPPT
MNSKKKASSQPGTKTKAESSDSSDSSDDEQPPKKAKIVPAKITASVPKPLAKKAETSS
DSESDSSSEDEKPSVKVAPKKAAAKPDAKSTPAAAAKKSAPAKKASSSSDSDSSSDEE
TTPKPSTKITPAKSAATPTSKTPTTSKATPTSKTLPAKQGTPKTSKKDSSSSDSSDSS
SDEDEKKTPAKPAAKTTPAKPAAKTTPAKPAAKTTPGKKAPTKKESSSSDSSDSSSED
EKKSSAKPGVKTTHGKAASKTDASAAAKSVPAKKASSSSDSDSDSSEEETTKTTKPLN
KPSPAVKSLPSKKAESSSDSSSSDSDSKKKTKPAKPPAKATTPVTPKVQVQNKKPSKA
NSTDSSSEEEGKSKQTTGKSPAAKATVPPKKSPVAVNKAKASSSSSDSSSDDEKQKP
KQPPAAKDVKQGAKATKPAPKKAADSSSEDSSSEDDDLKATKTNTAISKSPVTPKAKP
APKKESSSESSDSEDEKQKGKNISTPKTANSTPKAAAESSEESSSSEDEEKANGTSGK
RKRGESTGNAECETMTPENKKLKAKTPNTFPKVNKKEHNKNTPFRRVREEEVEINPLM
TDNAFEAKKGAQGDWGEKANNVLKFTKGKSFRHEKTKKKRGSYCGGAISTSVNSIKFD
SD"

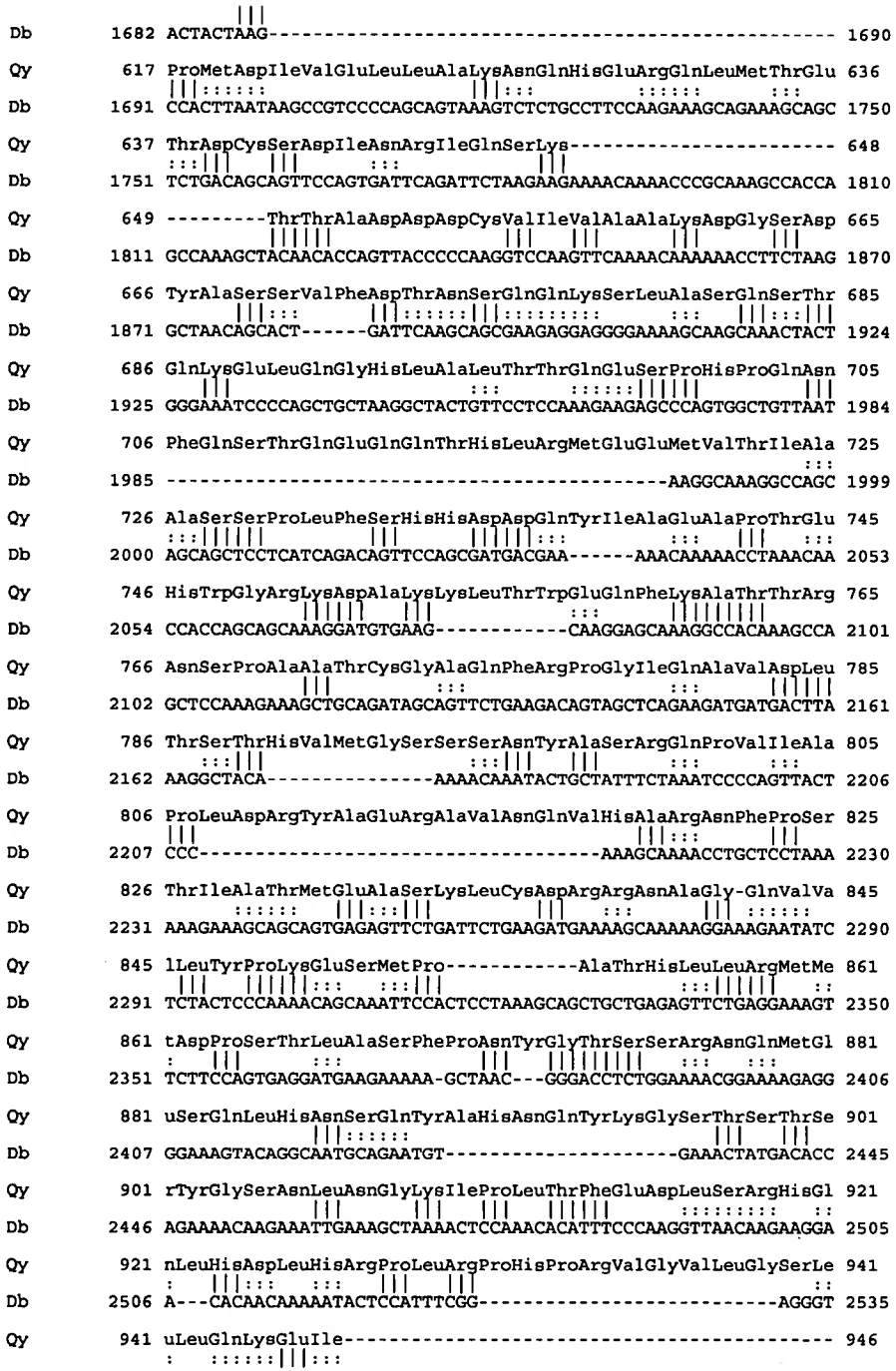
ORIGIN

Alignment Scores:

Pred. No.:	0.000499	Length:	3154
Score:	246.00	Matches:	212
Percent Similarity:	33.9%	Conservative:	156
Best Local Similarity:	19.6%	Mismatches:	428
Query Match:	4.5%	Indels:	289
DB:	5	Gaps:	40

US-09-828-068-2 (1-1057) x BC064870 (1-3154)

```
Qy         71 LysAspProLysPheCys------SerLeuSerArgIlePheHisAspGlnLysLysCys 88
Db        110 AAAGACAATGATTCCTGTCCAACATCTCTTCTGAATATATTTAGCGACTGGATGAAGTCA 169

Qy         89 AspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCys 108
Db        170 CCAGATGCAAAG---------------------------------------------------- 181

Qy        109 SerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
Db        182 --------------AAGAGAAAGCCATCAGCTAATGGT---TTGCCTAAGAAAAAGCCT 223

Qy        129 AlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheVal 148
Db        224 ACAAACGAGTCTTCCAGCAGCGAAGACAGCTCC---------------------------- 256

Qy        149 ProAlaSerValGlySerGlnLysValSerPro--------------SerThrGlnSer 163
Db        257 --------------AGTGAGGAGGATGAACCTCCTGCAAAAAAAATAGCAACTCAGTCT 301

Qy        164 SerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGlyAsnAsp 183
Db        302 GCTGAAGGGAAG-----------------------AAGCCTGTTGTGAAAGGCGTCCAA 337

Qy        184 SerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerPro 203
Db        338 GCAAAAAAAGCAAAGAGCAGCAGCGAGGACTCCAGTGATGAATCTGATTCTGAAAAAGAA 397

Qy        204 MetLys------AspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAla--------- 218
Db        398 ACAAAGAAACCTCCTGTTAAAAAACCCGCACAAACACCAAAGGGAGCAGCTTTAAAAACA 457

Qy        219 ----------------------AsnValSerGluAspAsnThrSerValAspValGly 230
Db        458 CCAACTCAAAAGAAGGCTGAGAGCTCCAGTTCTGAATCCAGCAGCTCAGAAGATGAAGCC 517

Qy        231 AlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnPro 250
Db        518 TCTAAGAAGAAGAAGCAATCTGTGACT------------------------------- 544

Qy        251 ProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLys 270
Db        545 ---AAAGTCCCTCCAAAGCAGACAGTAGTAAAGGCTGGTTTAGCAAACGACAATCAAAAA 601

Qy        271 ThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSer 290
Db        602 GCAGCAGACTCAAGTAGCAGTGAGGACTCTGAC------AGCCCCCCAGAAAAGAAGTCT 655

Qy        291 GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLys---AspProLysPro 309
Db        656 GCTGCTCCAAAGACACCTCCAGCCAAACCCGCCACAGCAGCTAAGCCACAAGCAAAGAAG 715

Qy        310 ValSerGlyGlnLys------CysGluGlnIleCysAsnGluProCysGluGluValVal 327
Db        716 ACAGCAGGGAAGAAGTCCAGTAGTGAGGACTCTTCAGACAGCTCAGATGAAGAGGCA--- 772

Qy        328 LeuLysArgSerSerLysSerLysArgLys------------------------------ 337
Db        773 ------AAGCCTGCAAAAAGCAAACCAAAGCCAGGTGTCTATAGTGCTGTACCACCACCT 826

Qy        338 ------ThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGln 355
Db        827 ACTATGAATTCTAAGAAAAAGGCTTCTTCACAGCCTGGCACCAAAACTAAGGCGGAAAGC 886

Qy        356 AlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSer 375
Db        887 AGTGACTCTTCAGACAGCAGTGATGATGAACAGCCACCAAAAAAAGCTAAAATTGTTCCA 946

Qy        376 GluIleIleAsnAlaAsn------------------------------GlnValGluAsp 385
Db        947 GCAAAGATTACTGCATCTGTCCCAAAACCCTTGGCCAAGAAAGCAGAGACCAGCTCAGAC 1006

Qy        386 SerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSer 405
Db       1007 AGTGAATCTGAT-----------------------TCCAGCAGTGAAGATGAAAAGCCA 1042

Qy        406 ThrIleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGly 425
Db       1043 AGTGTAAAGGTTGCCCCTAAAAAGGCTGCTGCAAAGCCAGATGCTAAAAGCACTCCAGCA 1102

Qy        426 GluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAsp 445
Db       1103 GCTGCTGCTAAAAAGTCTGCCCCAGCAAAGAAGGCTTCCAGCTCTTCAGAC------TCT 1156

Qy        446 AspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHis 465
Db       1157 GACTCTAGCAGT--------------GATGAAGAGACTACACCTAAACCTTCCACAAAA 1201

Qy        466 HisThrValAlaHisProAlaGlyAsnLeuSerAsnLys---------------LysVal 480
Db       1202 ATTACTCCTGCAAAGAGTGCTGCTACACCAACTTCCAAAACACCCACAACTAGCAAAGCA 1261

Qy        481 ThrProThrAlaSerThr---------------------------------------- 486
Db       1262 ACACCAACATCAAAGACCTTGCCTGCAAAACAAGGAACCCCCAAAACTTCCAAAAAGGAT 1321

Qy        487 ---GlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsn------------ 501
Db       1322 TCAAGCTCCTCAGACAGCAGTGACTCAAGCAGTGATGAAGATGAGAAGAAAACCCCTGCC 1381

Qy        502 ------MetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCys 519
Db       1382 AAACCTGCTGCTAAAACCACTCCTGCCAAGCCTGCTGCTAAAACCACTCCTGCCAAGCCT 1441

Qy        520 SerSerLys---------GlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAla 536
Db       1442 GCTGCTAAAACCACCCCAGGCAAGAAAGCTCCAACTAAGAAAGAATCCAGTTCTTCAGAT 1501

Qy        537 AlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSer 556
Db       1502 AGCTCAGATTCCAGCAGTGAAGATGAAAAGAAATCCTCTGCAAAGCCAGGAGTTAAGACT 1561

Qy        557 AlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysVal 576
Db       1562 ACACATGGGAAAGCTGCTTCCAAGACAGATGCTTCTGCTGCTGCAAAGTCTGTTCCTGCA 1621

Qy        577 SerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLys 596
Db       1622 AAGAAGGCTTCCAGCTCTTCAGACTCTGACTCTGATAGCTCAGAAGAGGAGACCACGAAA 1681

Qy        597 LysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIle 616
```



Search completed: March 7, 2006, 21:55:49
Job time : 12159 secs

GenCore version 5.1.7

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2006, 13:20:24 ; Search time 1439 Seconds
(without alignments)
4895.469 Million cell updates/sec

Title: US-09-828-068-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTNCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09828068/runat_06032006_163152_25112/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss03h
-USER=US09828068_@CGN_1_1_727_@runat_06032006_163152_25112 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5526	100.0	3896	8	ABQ77908	Abq77908 Rice OsEM
2	5526	100.0	4310	6	ABA93084	Aba93084 Oryza sat
3	5191	93.9	9455	6	ABA93085	Aba93085 Oryza sat
4	491.5	8.9	871	13	ADX09467	Adx09467 Plant ful
5	393	7.1	3291	8	ABZ42128	Abz42128 Arabidops
6	317	5.7	3679	13	ADT14740	Adt14740 Plant cDN
7	298	5.4	8646	3	AAA14873	Aaa14873 Nucleotid
8	298	5.4	8648	6	AAL43177	Aal43177 Arabidops
9	298	5.4	8648	8	ABQ77909	Abq77909 Arabidops
10	298	5.4	17341	3	AAA14872	Aaa14872 Genomic D
11	298	5.4	17341	6	AAL43176	Aal43176 Arabidops
12	241.5	4.4	7036	14	ADZ61911	Adz61911 Murine Nk
13	223.5	4.0	5286	5	AAS73156	Aas73156 DNA encod
14	223	4.0	2612	12	ACH89376	Ach89376 Human gen
15	223	4.0	4115	10	ADI21446	Adi21446 Novel hum
16	223	4.0	4327	14	ADY34541	Ady34541 DPCR1 (13
17	223	4.0	5238	10	ADI21918	Adi21918 Novel hum
18	223	4.0	10330	12	ADK60450	Adk60450 Angiogene
19	223	4.0	10330	12	ADK60751	Adk60751 Angiogene
20	223	4.0	10330	12	ADP73073	Adp73073 Angiogene
21	223	4.0	10330	14	ADY16535	Ady16535 DNA encod
22	223	4.0	10452	6	ABL68623	Abl68623 Kidney ca
23	223	4.0	10452	6	ABK84422	Abk84422 Human cDN
24	223	4.0	10452	12	ADK60475	Adk60475 Angiogene
25	223	4.0	10452	12	ADK60776	Adk60776 Angiogene
26	223	4.0	10452	12	ADP73098	Adp73098 Angiogene
27	223	4.0	11167	10	ADE53841	Ade53841 Human pro
28	222.5	4.0	279	6	ABL73385	Abl73385 Corn tass
29	221	4.0	5954	5	AAS80591	Aas80591 DNA encod
c 30	221	4.0	6143	5	AAS83843	Aas83843 DNA encod
31	221	4.0	6542	12	ADN04701	Adn04701 Antipsori
32	221	4.0	6542	12	ADQ19070	Adq19070 Human sof
33	221	4.0	7479	12	ADQ89533	Adq89533 cDNA enco
c 34	220	4.0	3953	4	ABA66656	Aba66656 Human foe
c 35	220	4.0	3953	4	ABA48746	Aba48746 Human bre
c 36	220	4.0	3953	4	ABA33721	Aba33721 Probe #12
c 37	220	4.0	3953	4	AAK40813	Aak40813 Human bon
c 38	220	4.0	3953	4	AAK15085	Aak15085 Human bra
c 39	220	4.0	3953	4	ABS40392	Abs40392 Human liv
c 40	220	4.0	3953	5	AAI07277	Aai07277 Probe #72
c 41	220	4.0	3953	6	ABS14767	Abs14767 Human gen
c 42	220	4.0	3953	12	ACH87407	Ach87407 Human gen
43	220	4.0	4383	11	ADN39109	Adn39109 Cancer/an
44	219	4.0	6201	8	ABX71059	Abx71059 Novel hum
45	217	3.9	3998	4	AAH29827	Aah29827 S cerevis

ALIGNMENTS

RESULT 1
ABQ77908
ID ABQ77908 standard; cDNA; 3896 BP.
XX
AC ABQ77908;
XX
DT 24-JAN-2003 (first entry)
XX
DE Rice OsEMF1-encoding cDNA, SEQ ID NO:1.
XX
KW Rice; OsEMF1; EMF; embryonic flower; plant; floral repressor;
KW reproductive development; flower development; transgenic plant;
KW antisense suppression; transgenic; reciprocal negative interaction;
KW flower meristem identity gene; flowering time; shoot development;
KW seed yield; agriculture; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 241. .3414
FT /*tag= a
FT /product= "OsEMF1 protein"
XX
PN WO200280659-A1.
XX
PD 17-OCT-2002.
XX
PF 05-APR-2002; 2002WO-US012675.

XX
PR 06-APR-2001; 2001US-00828068.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Moon Y, Chen L, Sung ZR;
XX
DR WPI; 2003-046831/04.
DR P-PSDB; ABB99878.
XX
PT New isolated OsEMF1 nucleic acid molecule and encoded polypeptide, useful
PT for plant genetic engineering, in particular controlling reproductive
PT development in rice.
XX
PS Claim 3; Fig 1; 47pp; English.
XX
CC The invention relates to a rice EMF (embryonic flower) gene designated
CC OsEMF1 (cDNA given in ABQ77908) and its encoded protein (ABB99878). The
CC OsEMF1 protein has a molecular weight of 116.4 kD and exhibits 37%
CC homology and 20% identity with Arabidopsis thaliana EMF1 protein
CC (ABB99879). OsEMF1, like other EMF gene products, acts as a floral
CC repressor, suppressing the transition from vegetative growth to
CC reproductive development. It also delays the inflorescence to flower
CC transition, indicating that there is a reciprocal negative interaction
CC between OsEMF and flower meristem identity genes. The invention also
CC encompasses a transgenic plant comprising an expression cassette
CC containing an OsEMF1 nucleic acid (particularly in an antisense
CC orientation) under the control of a plant promoter. OsEMF1 nucleic acids
CC and proteins may be used for controlling reproductive development in
CC plants, particularly monocotyledonous plants and especially rice. In
CC particular, OsEMF1 nucleic acids and proteins may be used to control
CC flowering time, shoot development and seed yield. Controlling or
CC inhibiting the expression of genes which mediate these processes enables
CC new varieties of rice with different flowering times and seed yield to be
CC developed. The present sequence represents OsEMF1 cDNA. Note: The present
CC sequence is described as SEQ ID NO:1 in the claims and sequence listing,
CC but the sequence referred to as SEQ ID NO:1 in the examples (not shown in
CC the specification) is described as an Arabidopsis thaliana EMF1 genomic
CC clone
XX
SQ Sequence 3896 BP; 1176 A; 791 C; 935 G; 994 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3896
Score:	5526.00	Matches:	1057
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-09-828-068-2 (1-1057) x ABQ77908 (1-3896)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840

Qy        201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900

Qy        221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960

Qy        241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020

Qy        261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080

Qy        281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140

Qy        301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440

Qy        401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680
Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
Db       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740
Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
Db       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800
Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
Db       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860
Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
Db       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920
Qy        561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
Db       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980
Qy        581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
Db       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040
Qy        601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100
Qy        621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160
Qy        641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220
Qy        661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280
Qy        681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340
Qy        701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
Db       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
Qy        721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
Db       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460
Qy        741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
Db       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520
Qy        761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
Db       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580
Qy        781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
Db       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640
Qy        801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
Db       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700
Qy        821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
Db       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760

Qy        841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
Db       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820
Qy        861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
Db       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880
Qy        881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
Db       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940
Qy        901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
Db       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000
Qy        921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
Db       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060
Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
Db       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120
Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
Db       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180
Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
Db       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240
Qy       1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
Db       3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300
Qy       1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
Db       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360
Qy       1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
Db       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3411
```

RESULT 2
ABA93084
ID ABA93084 standard; cDNA; 4310 BP.
XX
AC ABA93084;
XX
DT 11-APR-2002 (first entry)
XX
DE Oryza sativa brassinosteroid response related protein encoding cDNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
KW signal transfer system; brassinosteroid hormone; growth promotion;
KW increased yield; quality improvement; ripeness promotion;
KW stress relaxation; chemical resistance; gene; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 655. .3828
FT /*tag= a
FT /product= "brassinosteroid response related protein"
XX
PN JP2001327287-A.
XX
PD 27-NOV-2001.
XX
PF 19-MAY-2000; 2000JP-00149106.
XX

PR 19-MAY-2000; 2000JP-00149106.
XX
PA (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
PA (SEIB-) SEIBUTSUKEI TOKUTEI SANGYO GIJUTSU.
XX
DR WPI; 2002-135739/18.
DR P-PSDB; ABB05428.
XX
PT New gene involved in brassinosteroid responses useful for controlling the
PT effects such as growth promotion, increased yield, quality improvement,
PT ripeness promotion, stress relaxation and chemical resistance.
XX
PS Disclosure; Page 6-10; 19pp; Japanese.
XX
CC The present invention describes a polynucleotide encoding a plant gene
CC which can control the signal transfer system of brassinosteroid hormone.
CC The polynucleotide can be used for controlling the effects such as growth
CC promotion, increased yield, quality improvement, ripeness promotion,
CC stress relaxation and chemical resistance. The present sequence encodes a
CC brassinosteroid response related protein isolated from rice (Oryza
CC sativa), from the present invention
XX
SQ Sequence 4310 BP; 1237 A; 986 C; 1022 G; 1065 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	4310
Score:	5526.00	Matches:	1057
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-09-828-068-2 (1-1057) x ABA93084 (1-4310)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 714

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 774

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 834

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 894

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 954

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1014

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1074

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1134

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1194

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1254

Qy        201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1314

Qy        221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1374

Qy        241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1434

Qy        261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1494

Qy        281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1554

Qy        301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1614

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1674

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1734

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1735 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1794

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1795 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1854

Qy        401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1914

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1974

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2034

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2094

Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2154

Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2214

Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2274

Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2334
```

```
Qy       561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
Db      2335 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2394

Qy       581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
Db      2395 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2454

Qy       601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db      2455 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2514

Qy       621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db      2515 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2574

Qy       641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db      2575 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2634

Qy       661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db      2635 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2694

Qy       681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db      2695 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2754

Qy       701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
Db      2755 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2814

Qy       721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
Db      2815 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2874

Qy       741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
Db      2875 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2934

Qy       761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
Db      2935 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2994

Qy       781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
Db      2995 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3054

Qy       801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
Db      3055 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3114

Qy       821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
Db      3115 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3174

Qy       841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
Db      3175 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3234

Qy       861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
Db      3235 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3294

Qy       881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
Db      3295 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3354

Qy       901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
Db      3355 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3414

Qy       921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
Db      3415 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3474

Qy       941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
Db      3475 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3534

Qy       961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
Db      3535 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3594

Qy       981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
Db      3595 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3654

Qy      1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
Db      3655 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3714

Qy      1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
Db      3715 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3774

Qy      1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
Db      3775 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3825
```

RESULT 3
ABA93085
ID ABA93085 standard; DNA; 9455 BP.
XX
AC ABA93085;
XX
DT 11-APR-2002 (first entry)
XX
DE Oryza sativa brassinosteroid response related protein related DNA.
XX
KW Oryza sativa; rice; plant; brassinosteroid response; control;
KW signal transfer system; brassinosteroid hormone; growth promotion;
KW increased yield; quality improvement; ripeness promotion;
KW stress relaxation; chemical resistance; gene; ds.
XX
OS Oryza sativa.
XX
PN JP2001327287-A.
XX
PD 27-NOV-2001.
XX
PF 19-MAY-2000; 2000JP-00149106.
XX
PR 19-MAY-2000; 2000JP-00149106.
XX
PA (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
PA (SEIB-) SEIBUTSUKEI TOKUTEI SANGYO GIJUTSU.
XX
DR WPI; 2002-135739/18.
XX
PT New gene involved in brassinosteroid responses useful for controlling the
PT effects such as growth promotion, increased yield, quality improvement,
PT ripeness promotion, stress relaxation and chemical resistance.
XX
PS Disclosure; Page 13-16; 19pp; Japanese.
XX
CC The present invention describes a polynucleotide encoding a plant gene
CC which can control the signal transfer system of brassinosteroid hormone.
CC The polynucleotide can be used for controlling the effects such as growth
CC promotion, increased yield, quality improvement, ripeness promotion,
CC stress relaxation and chemical resistance. The present sequence
CC represents a DNA sequence related to a brassinosteroid response related
CC protein isolated from rice (Oryza sativa), from the present invention
XX
SQ Sequence 9455 BP; 2571 A; 2001 C; 2090 G; 2793 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	9455
Score:	5191.00	Matches:	1055
Percent Similarity:	68.1%	Conservative:	0
Best Local Similarity:	68.1%	Mismatches:	2
Query Match:	93.9%	Indels:	495
DB:	6	Gaps:	3

US-09-828-068-2 (1-1057) x ABA93085 (1-9455)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu  20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4223 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 4282

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4283 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 4342

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4343 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 4402

Qy         61 Ile---------------------------------------------------------- 61
Db       4403 AT-AAGGTAATCATTTTCTGTATTTCCAATTCCAGTATCGCGTTGTGGATGAATAATGAA 4461

Qy         61 ---------------------------------------------------------- 61
Db       4462 TCGGCATGTCATGCCATATTGCACTGTTTGATGGAAGAGTATGATTGATCGTGGTTTTTG 4521

Qy         61 ---------------------------------------------------------- 61
Db       4522 CACAGTTTGCTGTTGGGACTTATATGGTCATCTGTTTTGTACGATCGTATACACTGGGTC 4581

Qy         61 ---------------------------------------------------------- 61
Db       4582 GACATGCTTATGACTTTGGTTCGATTTAGGAAGTCAATACATCCACTACTAGCTCTATAT 4641

Qy         61 ---------------------------------------------------------- 61
Db       4642 CTAGCCATGTGAACTCATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAAAAAATATA 4701

Qy         61 ---------------------------------------------------------- 61
Db       4702 TATAATATTTGCATATATGTTGGTGTTTCATGTATCTTTATACTCTACGTACATCCATTA 4761

Qy         61 ---------------------------------------------------------- 61
Db       4762 ATATCTTCAATGTATGAATCTGAGCACATGATTGTGAGTGCTACACATATGCATGTCTGT 4821

Qy         61 ---------------------------------------------------------- 61
Db       4822 ATGTGTGTTCATTAGGTGTTTGATCATATTTGTTTGTGTTGGGGTGCGCATGCATTTATT 4881

Qy         61 ---------------------------------------------------------- 61
Db       4882 CAGGCCATGCTGTAGGCTGTAGCTAGATATTTGTGTTTGTATATTATTTCTGTTGAACAA 4941

Qy         61 ---------------------------------------------------------- 61
Db       4942 GCTGATTACTAATGAAATTAACCTTTTTGGGGTACACTCATATATTGGGCCCTACATTTT 5001

Qy         61 ---------------------------------------------------------- 61
Db       5002 TGTAATCATTTTTCCTTTGTGCTGAGGTTCAGCATAAAACTTTTTTATCATAAGCATGTT 5061

Qy         61 ---------------------------------------------------------- 61
Db       5062 TACATCCTAGGAGATTCTTAGAACTGATGGTTTCTTCATATTTGCATTATGTTTGATTTG 5121

Qy         61 ---------------------------------------------------------- 61
Db       5122 ATAGTCCATTATTATTTTAAGCCTTTTCAATTGTTTAGAGATTCTAGAGATGATATATAT 5181

Qy         61 ---------------------------------------------------------- 61
Db       5182 CAACCATAGACTTGTCACGTTTTGGTTTAATACTTTCTAGAACTAATTAGATTATTATTT 5241

Qy         61 ---------------------------------------------------------- 61
Db       5242 TTGTAGTTTATCCTGTCATGCTATTTGTATTATCTTTGAATTCAAACTGCAATACTTAGA 5301

Qy         61 ---------------------------------------------------------- 61
Db       5302 TTATCTTGAAGGTCCTCTTTTTCTGGACTGTACAAGCTATGTATGAAATGCCTACCTCCC 5361

Qy         61 ---------------------------------------------------------- 61
Db       5362 AGCATCCTTTAGATTATGTAGGGCCTTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC 5421

Qy         61 ---------------------------------------------------------- 61
Db       5422 ACGCAATGTGCTATATATATGTGCCATGCATCTTTATAATGATAATCTTATTTTTCTTGT 5481

Qy         62 ---ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg  80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5482 ACCAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 5541

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5542 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 5601

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5602 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 5661

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5662 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 5721

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5722 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 5781

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5782 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 5841

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5842 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 5901

Qy        201 AspSerProMetLys------------------------------------------- 205
              ||||||||||||
Db       5902 GATTCACCAATGAA-AGGTATGGTAGATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTT 5960

Qy        206 -----------------------------------------AspLeuGlnGlyProAla 211
                                                        ||||||||||||||||||
Db       5961 ATTTAAGGTATAGAATAAACTAATGTTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC 6020

Qy        212 GlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAla 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6021 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 6080

Qy        232 LeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6081 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 6140

Qy        252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6141 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 6200

Qy        272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6201 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGA 6260
```

```
Qy       292 LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
Db      6261 AAGGAACGTGATCAGGTTGCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCT 6320

Qy       312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
Db      6321 GGGCAGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGC 6380

Qy       332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
Db      6381 TCCAAATCTAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAA 6440

Qy       352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
Db      6441 CGCACTGCCCAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTG 6500

Qy       372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
Db      6501 CGGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTT 6560

Qy       392 HisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMet 411
Db      6561 CATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATG 6620

Qy       412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
Db      6621 GAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCA 6680

Qy       432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
Db      6681 AGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATG 6740

Qy       452 AsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisPro 471
Db      6741 AACTGGCTGAATGGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCA 6800

Qy       472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
Db      6801 GCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAG 6860

Qy       492 AsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVal 511
Db      6861 AATGATACTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTGTCAGCATGTA 6920

Qy       512 SerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGly 531
Db      6921 TCAGAAATCTCCACACAGAGGTGCTCATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGG 6980

Qy       532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
Db      6981 AAAACACATTCAGCTGCTAGTACCAAATATGGTGGTGAAAGCACCAGAAATGGTCAGAAC 7040

Qy       552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
Db      7041 ATACATGTACTCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGT 7100

Qy       572 HisSerAlaLys---------------------------------------------- 575
Db      7101 CACTCGGCAAAGGTACGAATTTTGTGAATCATGAGGAATTTTTGCTTTTTAAATTGACTG 7160

Qy       575 ---------------------------------------------------------- 575
Db      7161 AATCAACATTTATCTGTATGAAGGAATAATATTGGTGCATAACAATGTTAAGAAATATGC 7220

Qy       575 ---------------------------------------------------------- 575
Db      7221 ATACAATGTTTATTTATATGCTTTCCACTGTTCTTCTTTACTTATGTTTTGATACTCTTT 7280

Qy       575 ---------------------------------------------------------- 575
Db      7281 TTGTGTGTGCGTGCATGTGTGCATGTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGC 7340

Qy       575 ---------------------------------------------------------- 575
```

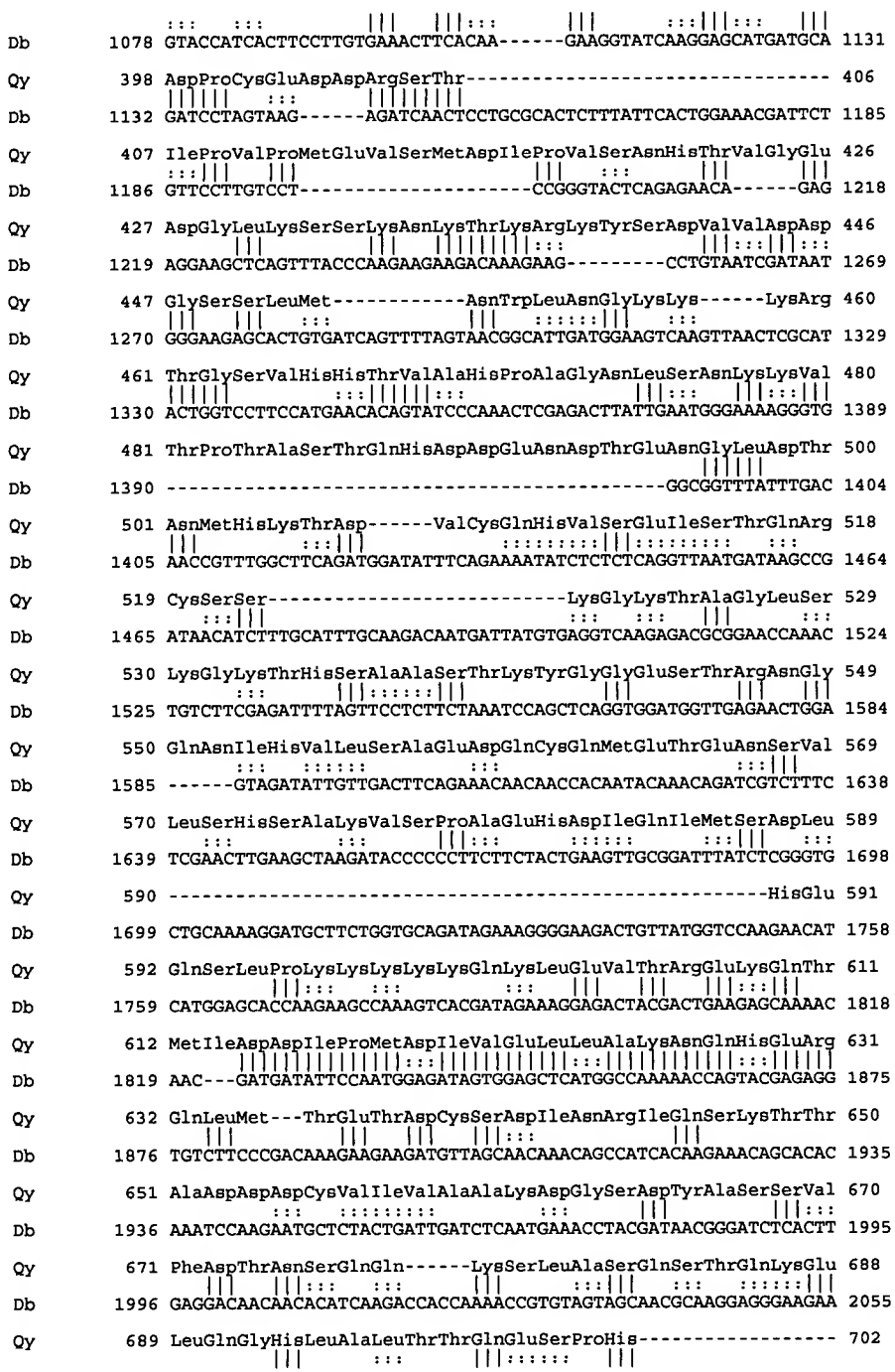
```
Db      7341 ACGTGCGTGGCGCAATATTCTTTTTTAGACTCATATTATAGTGATTGTAATGGACTGACA 7400

Qy       576 ----------------------ValSerProAlaGluHisAspIleGlnIleMetSerAs 588
Db      7401 TTTTCCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGA 7460

Qy       588 pLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGl 608
Db      7461 CCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGA 7520

Qy       608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGl 628
Db      7521 AAAACAGACCATGATAGATGACATCCCCATGGATATTGTTGAACTGCTAGCTAAAAACCA 7580

Qy       628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLy 648
Db      7581 GCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCTGACATCAACCGTATTCAATCCAA 7640

Qy       648 sThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe 668
Db      7641 GACAACTGCTGATGATGATTGTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCATC 7700

Qy       668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGl 688
Db      7701 AAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGA 7760

Qy       688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSe 708
Db      7761 GTTACAGGGTCATTTGGCATTGACCACACAAGAGTCTCCACATCCTCAGAACTTTCAGTC 7820

Qy       708 rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
Db      7821 TACTCAGGAACAGCAGACACATTTGCGGATGGAAGAAATGGTCACTATTGCTGCAAGCTC 7880

Qy       728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGl 748
Db      7881 ACCACTATTTTCACATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGG 7940

Qy       748 yArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
Db      7941 CCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCC 8000

Qy       768 oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh 788
Db      8001 AGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTAC 8060

Qy       788 rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs 808
Db      8061 TCATGTCATGGGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGA 8120

Qy       808 pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl 828
Db      8121 CCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGC 8180

Qy       828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
Db      8181 AACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCC 8240

Qy       848 oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe 868
Db      8241 TAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATTAGCAAG 8300

Qy       868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888
Db      8301 CTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCA 8360

Qy       888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl 908
Db      8361 GTATGCACATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAATGG 8420

Qy       908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
```

```
Db       8421 AAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACAGACC 8480

Qy        928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8481 TTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAA 8540

Qy        948 nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleTh 968
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8541 CTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAAC 8600

Qy        968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8601 ATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGC 8660

Qy        988 aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr 1008
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8661 AAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAG 8720

Qy       1008 gAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8721 GAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATCG 8780

Qy       1028 gPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSe 1048
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8781 GTTTGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATTTTACTACAATCAG 8840

Qy       1048 rAsnAspAsnGluTyrMetAspTyrArg 1057
              |||||||||||||||||||||||||||
Db       8841 TAACGATAACGAGTATATGGATTACCGC 8868
```

```
RESULT 4
ADX09467
ID   ADX09467 standard; cDNA; 871 BP.
XX
AC   ADX09467;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 4042.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 4042; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 871 BP; 225 A; 197 C; 230 G; 219 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                2.9e-27        Length:          871
Score:                    491.50         Matches:         107
Percent Similarity:       65.8%          Conservative:    18
Best Local Similarity:    56.3%          Mismatches:      50
Query Match:              8.9%           Indels:          15
DB:                       13             Gaps:            6

US-09-828-068-2 (1-1057) x ADX09467 (1-871)

Qy        871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu------SerGlnLeuHisAsnSerGln 888
              ||||||      :::::::|||   |||||||||      :::|||         ||||||
Db          2 AACTACCAAAGAGCTAATAGGCGGCAGATGGAGCTTCAAACGCAAGCTCTCGGCTCGCAG 61

Qy        889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907
              |||       |||       |||       |||   ||||||||||||||||||
Db         62 TACACACAGCATGAGCATTACAATGCGTCACCGAGCACGTCATATGGAAGCCAA------ 115

Qy        908 GlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926
                  |||||||||   ||||||||||||   ||||||          ||||||
Db        116 ---------CCGCTGACACTGGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGCGC 166

Qy        927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        167 AGGCCTTTACGCCCTCACCCTCGTGTCGGTGTGCTCGGTTCATTGCTGCAGCAGGAGATC 226

Qy        947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGly 966
              ||||||||||||   |||||||||   ||||||||||||:::||||||          |||
Db        227 GCAAACTGGTCTGGGAACTGCGGGCCGCAGTCTGGGTACAGACTAGGTGATTGTAAAGGG 286

Qy        967 IleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPhe 986
                          ||||||      :::::|||   |||:::
Db        287 ACAAGGCCGCCTGATGCCAACAGAGCAGGAAACTACGAGACCTTGAGC------------ 334

Qy        987 SerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla------AspPhe 1004
              ||||||   ||||||   ||||||||||||||||||||:::|||      ::::::
Db        335 TCGGCAGGATGGAACACCCTGCAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGTAC 394

Qy       1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024
                              |||       |||   |||   |||:::   :::|||
Db        395 CGGTTTCCATGGTATGGTACGGGTCAGCCTTCGACCAGTGGCAATGGGAGGACCATTCAT 454

Qy       1025 ProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPhe 1044
              |||||||||:::   ||||||:::|||||||||:::||||||:::|||||||||||||||
Db        455 CCGTTGGATAAGCTCGTGAGGAAGGATATCTGTGTGACTAACAGAAACCCAGCCGATTTC 514

Qy       1045 ThrThrIleSerAsnAspAsnGluTyrMet 1054
```

```
                ||| ||||||::: ||||||||||||
Db        515 ACTGTAATTAGTGACAAGAACGAGTACATG 544

RESULT 5
ABZ42128
ID   ABZ42128 standard; cDNA; 3291 BP.
XX
AC   ABZ42128;
XX
DT   27-FEB-2003  (first entry)
XX
DE   Arabidopsis thaliana gene #112 modulated by PTGS.
XX
KW   Posttranscriptional gene silencing; PTGS; plant; transformation; gene;
KW   ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3291
FT                   /*tag=  a
XX
PN   WO200281695-A2.
XX
PD   17-OCT-2002.
XX
PF   05-APR-2002; 2002WO-EP003806.
XX
PR   06-APR-2001; 2001US-0282049P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA   (FRIE-) FRIEDRICH MIESCHER INST.
XX
PI   Zhu T,  Glazov EA,  Meins F,  Wang X,  Chang H;
XX
DR   WPI; 2003-103337/09.
DR   P-PSDB; ABP81284.
XX
PT   Novel polynucleic acid segment useful for modulating gene expression
PT   within a cell by posttranscriptional gene silencing, and for augmenting a
PT   plant cell genome.
XX
PS   Claim 18; Page 403-405; 438pp; English.
XX
CC   The invention relates to a novel isolated polynucleic acid segment
CC   modulated within a cell by posttranscriptional gene silencing (PTGS). The
CC   invention specifically relates to a method to identify an expression
CC   product that is modulated by PTGS. The polynucleotide is useful for
CC   modulating the gene expression within a cell by PTGS, by introducing the
CC   polynucleic acid into a cell and expressing the nucleic acid segment in
CC   the cell to form a product. The polynucleic acid segment is also useful
CC   for augmenting a cell genome, and for augmenting a plant genome, by
CC   contacting a plant cell with the segment to produce a transformed plant
CC   cell, and growing the transformed plant cell to produce a differentiated
CC   transformed plant. The sequences shown in ABZ42017 - ABZ42142 represent
CC   segments of A. thaliana cDNA modulated by PTGS
XX
SQ   Sequence 3291 BP; 1066 A; 689 C; 727 G; 809 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.75e-19        Length:         3291
Score:                  393.00          Matches:        248
Percent Similarity:     36.8%           Conservative:   191
Best Local Similarity:  20.8%           Mismatches:     401
Query Match:            7.1%            Indels:         352
DB:                     8               Gaps:           58

US-09-828-068-2 (1-1057) x ABZ42128 (1-3291)

Qy         55 GlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLys 74
              :::|||:::|||||||||:::||||||:::||||||      ::::::::::|||   :::
Db         73 AAATGTGATCATTTTTCCATGCGTGGATTCGTAGCTGAAACTCGTGAGAGAGACCTTAGA 132

Qy         75 PheCys-----------------SerLeuSerArgIlePheHisAspGlnLysLysCys 88
                 |||                 |||:::|||   :::
Db        133 AAATGTTGGCCGTTTTCAGAAGAGAGTGTTAGTTTAGTA---------------------- 171

Qy         89 AspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCys 108
              |||:::         :::             ||||||   |||||||||   |||   |||
Db        172 GACCAACAAAGCTATACTCTTCCTACTTTATCTGTTCCAAAGTTTAGATGGTGGCATTGT 231

Qy        109 SerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
                 |||:::     :::             :::         |||:::
Db        232 ATGAGCTGCATCAAAGATATA-----------GATGCTCATGGGCCAAAAGATTGTGGA 279

Qy        129 AlaLysGlnAsnGlyThrSerAspGly---CysSerIleThrPheValArgSerThrPhe 147
                 |||     :::   |||     |||:::       :::|||   |||
Db        280 CTGCATTCAAACTCAAAAGCTATTGGAAACTCTTCTGTTATCGAAAGTAAAAGCAAGTTC 339

Qy        148 ValProAlaSerValGlySerGlnLysValSerProSerThr------------------ 161
                   ::::::       :::          |||
Db        340 AATTCGCTAACTATCATTGATCACGAGAAAGAAAAGAAAACTGATATTGCAGATAATGCT 399

Qy        162 ---GlnSerSerGlnGlyLysAsnAlaAspArg--------------SerThrLeuPro 175
                 :::         |||  |||  :::                 :::|||
Db        400 ATTGAGGAGAAAGTGGGTGTAAACTGTGAGAATGATGATCAGACAGCTACTACGTTTCTC 459

Qy        176 LysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAla 195
              |||     :::                         |||   |||
Db        460 AAGAAAGCACGT--------------------GGTCGACCTATGGGT------------ 486

Qy        196 AlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAsp 215
                 :::|||      |||      :::  |||      |||  |||
Db        487 ---GCTTCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG--------- 534

Qy        216 ValAlaAlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeu 232
              |||      |||  |||::::::                |||:::|||:::  :::
Db        535 GTAGGAAACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGG 594

Qy        233 ProGluValProGlnIleThrTrpHisIleGluValAsnGlyAla--------------- 247
                 |||     :::         :::         |||:::
Db        595 AAAGAGAAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGT 654

Qy        248 ---------AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsn 264
                      |||  ||||||          |||      :::  :::
Db        655 GTGGTTGAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATG 714

Qy        265 GluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLys 284
              |||  :::||||||   :::|||         :::|||
Db        715 GAATGCGATAACGGGTCATCAGAAAGTATAAATCTTGCT---------------------- 753

Qy        285 AspProAsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCys 301
                       ||||||||   :::|||            :::|||
Db        754 -----------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGT 801

Qy        302 AsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGlu 321
                 |||      :::  |||          |||             :::
Db        802 GAGTTGCTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------- 849

Qy        322 ProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLys 341
                 ||||||      ||||||:::  |||          ||||||      :::|||
Db        850 ------GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAG 897

Qy        342 LeuMetLysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp 360
              |||:::   :::      :::   |||:::   :::             |||:::
Db        898 TTGTTACCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAA 957

Qy        361 ------AlaLysLeuCys---------------------------------------- 364
                   :::|||   |||
Db        958 AATGCTTCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTT 1017

Qy        365 --------------------ArgArgLysProLysLysValArgLeuLeuSerGluIle 377
                                  :::::::::  ::::::  :::::::::  |||
Db       1018 GACAGAACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTT 1077

Qy        378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397
```


Db 1078 GTACCATCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCA 1131
Qy 398 AspProCysGluAspAspArgSerThr---------- 406
Db 1132 GATCCTAGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCT 1185
Qy 407 IleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGlu 426
Db 1186 GTTCCTTGTCCT----------CCGGGTACTCAGAGAACA------GAG 1218
Qy 427 AspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp 446
Db 1219 AGGAAGCTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAAT 1269
Qy 447 GlySerSerLeuMet----------AsnTrpLeuAsnGlyLysLys------LysArg 460
Db 1270 GGGAAGAGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCAT 1329
Qy 461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db 1330 ACTGGTCCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTG 1389
Qy 481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
Db 1390 ----------GGCGGTTTATTTGAC 1404
Qy 501 AsnMetHisLysThrAsp------ValCysGlnHisValSerGluIleSerThrGlnArg 518
Db 1405 AACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAATGATAAGCCG 1464
Qy 519 CysSerSer----------LysGlyLysThrAlaGlyLeuSer 529
Db 1465 ATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCGGAACCAAAC 1524
Qy 530 LysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGly 549
Db 1525 TGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTGAGAACTGGA 1584
Qy 550 GlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerVal 569
Db 1585 ------GTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCTTTC 1638
Qy 570 LeuSerHisSerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeu 589
Db 1639 TCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGGGTG 1698
Qy 590 ----------HisGlu 591
Db 1699 CTGCAAAAGGATGCTTCTGGTGCAGATAGAAAGGGGAAGACTGTTATGGTCCAAGAACAT 1758
Qy 592 GlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThr 611
Db 1759 CATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGGAGACTACGACTGAAGAGCAAAAC 1818
Qy 612 MetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArg 631
Db 1819 AAC---GATGATATTCCAATGGAGATAGTGGAGCTCATGGCCAAAAACCAGTACGAGAGG 1875
Qy 632 GlnLeuMet---ThrGluThrAspCysSerAspIleAsnArgIleGlnSerLysThrThr 650
Db 1876 TGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAACAGCCATCACAAGAAACAGCACAC 1935
Qy 651 AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVal 670
Db 1936 AAATCCAAGAATGCTCTACTGATTGATCTCAATGAAACCTACGATAACGGGATCTCACTT 1995
Qy 671 PheAspThrAsnSerGlnGln------LysSerLeuAlaSerGlnSerThrGlnLysGlu 688
Db 1996 GAGGACAACAACACATCAAGACCACCAAAACCGTGTAGTAGCAACGCAAGGAGGGAAGAA 2055
Qy 689 LeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHis---------- 702
Db 2056 ---------CATTTTCCTATGGGAAGACAGCAGAACTCTCATGACTTCTTCCCAATAAGT 2106
Qy 703 ----------ProGlnAsn---------PheGlnSerThrGlnGluGlnGlnThrHis 715
Db 2107 CAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTCCTACCCAAGAAAACCGA------ 2160
Qy 716 LeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAsp 735
Db 2161 ----------GCCAGCTCCATCCGGTTTTCTGGTCACAAC 2190
Qy 736 AspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeu 755
Db 2191 TGTCAGTGGCTTGGGAATTTGCCAACTGTG------GGTAATCAGAACCCTTCTCCATCC 2244
Qy 756 ThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAla--- 774
Db 2245 TCATTTCGGGTATTACGTGCT----------TGTGATACTTGCCAGAGTGTT 2286
Qy 775 ------GlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetGly 792
Db 2287 CCTAATCAATACAGA----------GAAGCTTCTCATCCAATTTGG 2322
Qy 793 SerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaGlu 812
Db 2323 CCATCTTCC----------ATGATACCACCACAGAGTCAATACAAGCCA 2361
Qy 813 ArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAla 832
Db 2362 GTTTCTTTAAATATTAATCAGTCAACAAAT---CCGGGTACGCTTTCA------CAGGCA 2412
Qy 833 SerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProLysGlu------ 850
Db 2413 TCTAACAATGAAAATACATGGAACCTCAACTTTGTTGCTGCCAACGGGAAGCAAAAATGT 2472
Qy 850 ---------- 850
Db 2473 GGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATGCTGCTGGGGTTAGTAGTAGTAGT 2532
Qy 851 ----------SerMetProAlaThrHisLeuLeuArg 859
Db 2533 AGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTATACCGGCATTGCATCTACTCAGC 2592
Qy 860 MetMetAspProSerThrLeuAlaSerPhePro------AsnTyrGlyThrSer------ 875
Db 2593 CTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTGACCAACACGGAAACACTAAATTT 2652
Qy 876 SerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLys 895
Db 2653 ACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAGAGTTTATAGAGCTTCAAACAGGG 2712
Qy 896 GlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGlu 915
Db 2713 GACTCTAGTAAGTCAGCCTACTCAACTAAG---------CAGATACCT------TTTGAT 2757
Qy 916 AspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeuArgPro---HisProArg 934
Db 2758 CTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGAAGAGTTTCCCCATCACTCCACCT 2817
Qy 935 ValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGly 954
Db 2818 ATTGGGACG------TCTTCACTTTCATTTCAAAATGCTTCATGGAGTCCTCAT------ 2865
Qy 955 ThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArg 974
Db 2866 ----------CATCAAGAGAAGAAA 2880
Qy 975 LysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGln 994
Db 2881 ACCAAGAGAAAAGACACC----------TTTGCTCCTGTTTACAATACTCATGAA 2925
Qy 995 LeuGlySerValSerSerSerAlaAsp----------PheLeuSerAlaArgAsn 1009
Db 2926 AAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGTTCCAGCTGCTGGGAGCATCGAAT 2985

```
Qy       1010 SerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeu--------- 1026
              |||                                  |||:::   ||||||
Db       2986 TCC--------------------------------ATGATGCTTCCTTTGAAATTTCAC 3012

Qy       1027 ------AspArgPheValArgGlnAsp---------------------------------- 1033
                    |||:::      :::|||
Db       3013 ATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAGAGAGCTGCAATAACAATGCCTCT 3072

Qy       1034 --------------------------------IleCysIleThrAsnLysAsnProAla 1042
                                              :::|||      |||:::|||||||||
Db       3073 GCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGTGCAGCGTCAATAGAAACCCTGCT 3132

Qy       1043 AspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
              |||||||||               |||   ||||||
Db       3133 GATTTCACCATTCCTGAACCTGGGAATGTTTACATG 3168
```

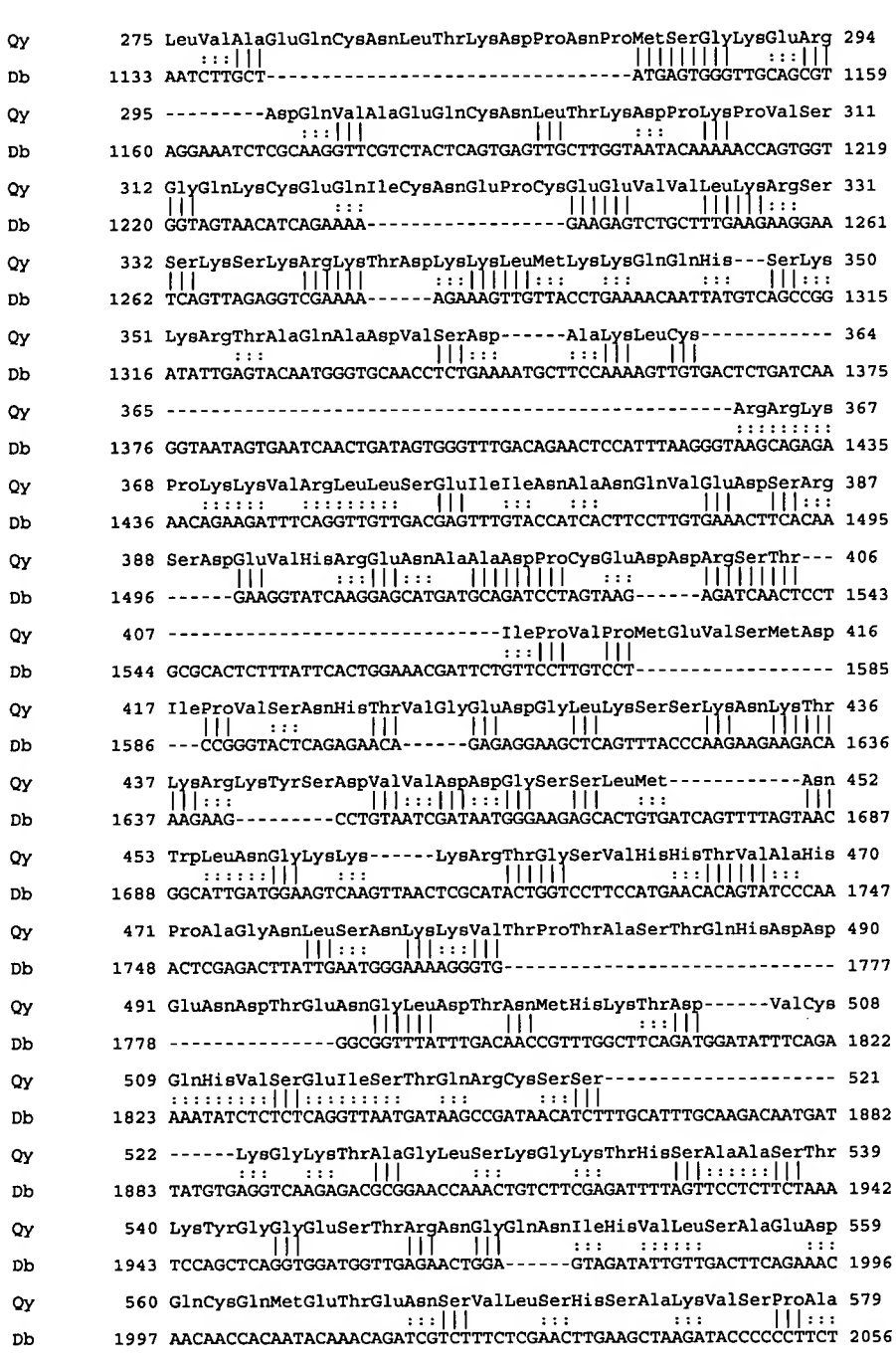
RESULT 6
ADT14740
ID ADT14740 standard; cDNA; 3679 BP.
XX
AC ADT14740;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant cDNA, Seq ID 66.
XX
KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactomannan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI Kovalic DK;
XX
DR WPI; 2004-757369/74.
XX
PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
PS Claim 1; SEQ ID NO 66; 14pp; English.
XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 3679 BP; 1151 A; 811 C; 822 G; 895 T; 0 U; 0 Other;

Alignment Scores:

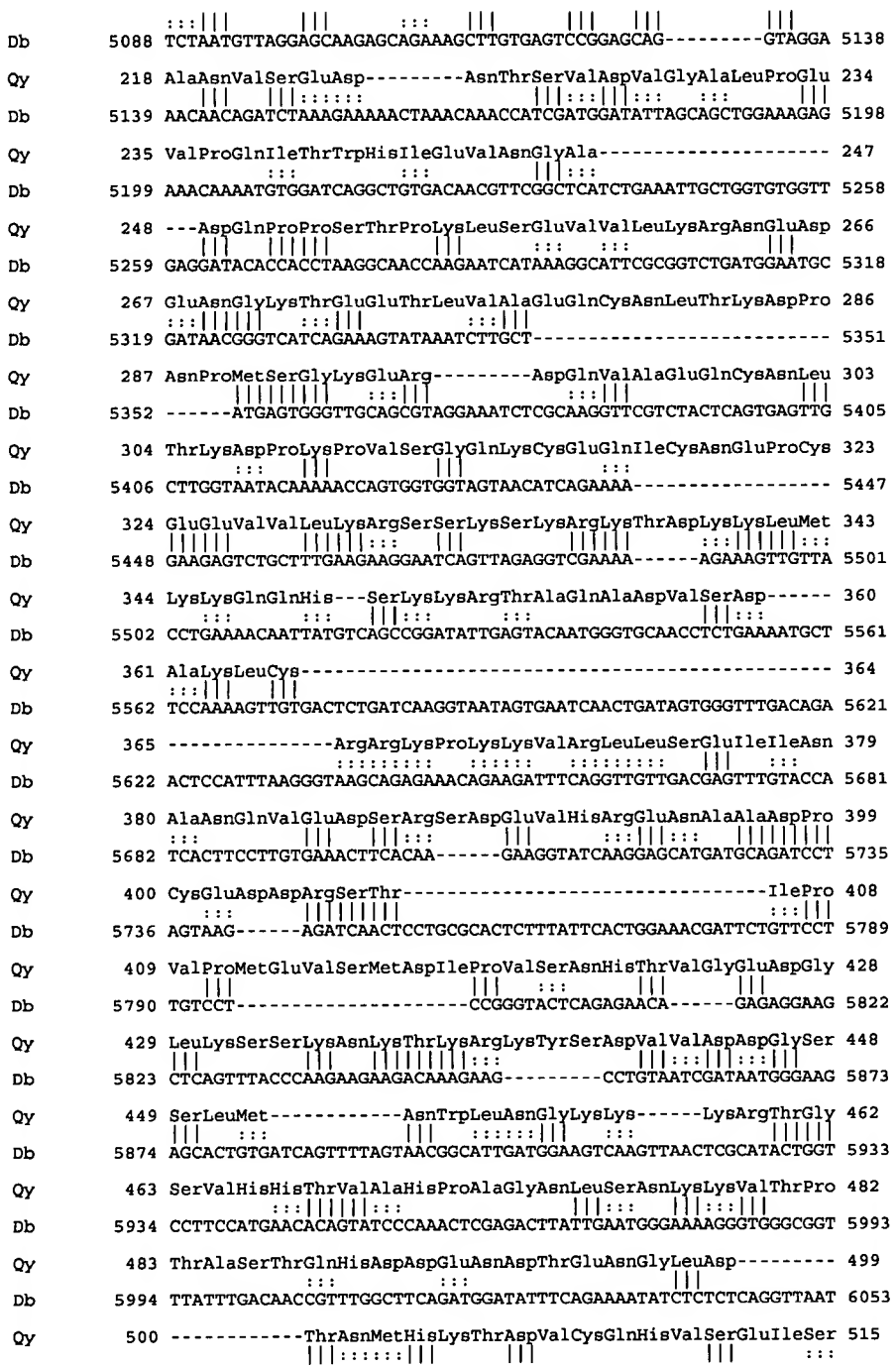
Pred. No.:	2.03e-13	Length:	3679
Score:	317.00	Matches:	233
Percent Similarity:	35.8%	Conservative:	176
Best Local Similarity:	20.4%	Mismatches:	391
Query Match:	5.7%	Indels:	342
DB:	13	Gaps:	54

US-09-828-068-2 (1-1057) x ADT14740 (1-3679)

```
Qy         83 HisAspGlnLysLysCysAspGluHis---------LysAlaSerSerSerProPheSer 99
              |||      |||  |||      |||         ||||||      :::
Db        647 CATGGGCCAAAAGATTGTGGACTGCATTCAAACTCAAAAGCTATTGGAAACTCTTCTGTT 706

Qy        100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
              :::          :::::::::          :::|||    |||
Db        707 ATCGAAAGTAAAAGCAAGTTCAATTCGCTAACTATCATTGATCACGAGAAAGAAAAGAAA 766

Qy        120 AsnGlyThrAlaProArgThrLeuProAlaLys------GlnAsnGlyThrSerAspGly 137
                   |||         :::      |||           |||     :::|||
Db        767 ACTGATATTGCAGATAATGCTATTGAGGAGAAAGTGGGTGTAAACTGTGAGAATGATGAT 826

Qy        138 CysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysVal 157
               :::   |||           ||||||:::      :::   |||
Db        827 CAGACAGCTACT-----------ACGTTTCTCAAGAAAGCACGTGGTCGACCTATGGGT 874

Qy        158 SerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSer 177
              :::   :::  :::|||          |||                |||
Db        875 GCTTCTAATGTTAGGAGCAAGAGCAGAAAG-----------------CTTGTGAGTCCG 916

Qy        178 ValGlnGluGlyAsnAsp---------SerLysCysAsnAlaProSerGlyLysAsnGly 194
                 |||   ||||||:::          |||   |||   ||||||
Db        917 GAGCAGGTAGGAAACAACAGATCTAAAGAAAAACTAAACAAACCATCG----------- 964

Qy        195 AlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyr 214
                                          |||:::          :::
Db        965 ----------------------------ATGGATATTAGCAGCTGGAAAGAGAAACAA 994

Qy        215 AspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGlu 234
              :::|||      |||:::      :::|||       |||   :::
Db        995 AATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT------ 1048

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrPro 254
                                                 |||   ||||||
Db       1049 -----------------------------------GAGGATACACCACCTAAGGCAACC 1072

Qy        255 LysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThr 274
              |||      :::  :::           |||  :::||||||   :::|||
Db       1073 AAGAATCATAAAGGCATTCGCGGTCTGATGGAATGCGATAACGGGTCATCAGAAAGTATA 1132
```


Qy 275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
Db 1133 AATCTTGCT----------ATGAGTGGGTTGCAGCGT 1159
Qy 295 ---------AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
Db 1160 AGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTGCTTGGTAATACAAAAACCAGTGGT 1219
Qy 312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
Db 1220 GGTAGTAACATCAGAAAA----------GAAGAGTCTGCTTTGAAGAAGGAA 1261
Qy 332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHis---SerLys 350
Db 1262 TCAGTTAGAGGTCGAAAA------AGAAAGTTGTTACCTGAAAACAATTATGTCAGCCGG 1315
Qy 351 LysArgThrAlaGlnAlaAspValSerAsp------AlaLysLeuCys---------- 364
Db 1316 ATATTGAGTACAATGGGTGCAACCTCTGAAAATGCTTCCAAAAGTTGTGACTCTGATCAA 1375
Qy 365 ----------ArgArgLys 367
Db 1376 GGTAATAGTGAATCAACTGATAGTGGGTTTGACAGAACTCCATTTAAGGGTAAGCAGAGA 1435
Qy 368 ProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg 387
Db 1436 AACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCATCACTTCCTTGTGAAACTTCACAA 1495
Qy 388 SerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThr--- 406
Db 1496 ------GAAGGTATCAAGGAGCATGATGCAGATCCTAGTAAG------AGATCAACTCCT 1543
Qy 407 ----------IleProValProMetGluValSerMetAsp 416
Db 1544 GCGCACTCTTTATTCACTGGAAACGATTCTGTTCCTTGTCCT---------- 1585
Qy 417 IleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThr 436
Db 1586 ---CCGGGTACTCAGAGAACA------GAGAGGAAGCTCAGTTTACCCAAGAAGAAGACA 1636
Qy 437 LysArgLysTyrSerAspValValAspAspGlySerSerLeuMet----------Asn 452
Db 1637 AAGAAG---------CCTGTAATCGATAATGGGAAGAGCACTGTGATCAGTTTTAGTAAC 1687
Qy 453 TrpLeuAsnGlyLysLys------LysArgThrGlySerValHisHisThrValAlaHis 470
Db 1688 GGCATTGATGGAAGTCAAGTTAACTCGCATACTGGTCCTTCCATGAACACAGTATCCCAA 1747
Qy 471 ProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAsp 490
Db 1748 ACTCGAGACTTATTGAATGGGAAAAGGGTG---------- 1777
Qy 491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAsp------ValCys 508
Db 1778 ----------GGCGGTTTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGA 1822
Qy 509 GlnHisValSerGluIleSerThrGlnArgCysSerSer---------- 521
Db 1823 AAATATCTCTCTCAGGTTAATGATAAGCCGATAACATCTTTGCATTTGCAAGACAATGAT 1882
Qy 522 ------LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThr 539
Db 1883 TATGTGAGGTCAAGAGACGCGGAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAA 1942
Qy 540 LysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAsp 559
Db 1943 TCCAGCTCAGGTGGATGGTTGAGAACTGGA------GTAGATATTGTTGACTTCAGAAAC 1996
Qy 560 GlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAla 579
Db 1997 AACAACCACAATACAAACAGATCGTCTTTCTCGAACTTGAAGCTAAGATACCCCCCTTCT 2056
Qy 580 GluHisAspIleGlnIleMetSerAspLeu---------- 589
Db 2057 TCTACTGAAGTTGCGGATTTATCTCGGGTGCTGCAAAAGGATGCTTCTGGTGCAGATAGA 2116
Qy 590 ----------HisGluGlnSerLeuProLysLysLysLysGln 601
Db 2117 AAGGGGAAGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGAT 2176
Qy 602 LysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleVal 621
Db 2177 AGAAAGGAGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTG 2233
Qy 622 GluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSer 640
Db 2234 GAGCTCATGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGC 2293
Qy 641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db 2294 AACAAACAGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTC 2353
Qy 661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------Lys 678
Db 2354 AATGAAACCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAA 2413
Qy 679 SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGln 698
Db 2414 CCGTGTAGTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAG 2464
Qy 699 GluSerProHis----------ProGlnAsn--------- 705
Db 2465 CAGAACTCTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATC 2524
Qy 706 PheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAla 725
Db 2525 TTTCCTCCTACCCAAGAAAACCGA---------- 2548
Qy 726 AlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGlu 745
Db 2549 GCCAGCTCCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG 2608
Qy 746 HisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db 2609 ------GGTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT--------- 2653
Qy 766 AsnSerProAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAla 782
Db 2654 ---------TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------- 2689
Qy 783 ValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnPro 802
Db 2690 ---------GAAGCTTCTCATCCAATTTGGCCATCTTCC---------- 2719
Qy 803 ValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsn 822
Db 2720 ATGATACCACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT 2779
Qy 823 PheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGly 842
Db 2780 ---CCGGGTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAAC 2830
Qy 843 GlnValValLeuTyrProLysGlu---------- 850
Db 2831 TTTGTTGCTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGC 2890
Qy 850 ---------- 850
Db 2891 AAACATGCTGCTGGGGTTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAG 2950
Qy 851 ---SerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhe 869
Db 2951 AGCTCTATACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACT 3010
Qy 870 Pro------AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHis 885

```
Db       3011 CCCGCTGACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAG 3070

Qy        886 AsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsn 905

Db       3071 TCCAAAGAGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG 3130

Qy        906 LeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeu 925

Db       3131 ---------CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCT 3175

Qy        926 HisArgProLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLys 944

Db       3176 TCCCGGAAGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTT 3229

Qy        945 GluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSer 964

Db       3230 CAAAATGCTTCATGGAGTCCTCAT---------------------------------- 3253

Qy        965 ThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGly 984

Db       3254 --------------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC------------ 3286

Qy        985 MetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp--- 1003

Db       3287 ---TTTGCTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAA 3343

Qy       1004 ------------PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLys 1019

Db       3344 GCGAAGTTCCAGCTGCTGGGAGCATCGAATTCC-------------------------- 3376

Qy       1020 GlyLysMetValHisProLeu---------------AspArgPheValArgGlnAsp--- 1033

Db       3377 ------ATGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGA 3430

Qy       1033 ------------------------------------------------------------ 1033

Db       3431 AAAGCAGAGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCC 3490

Qy       1034 ---IleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGlu 1052

Db       3491 ATTGTGTGCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTT 3550

Qy       1053 TyrMet 1054

Db       3551 TACATG 3556

RESULT 7
AAA14873
ID   AAA14873 standard; DNA; 8646 BP.
XX
AC   AAA14873;
XX
DT   08-AUG-2000  (first entry)
XX
DE   Nucleotide sequence of the EMF-1 gene and its promoter.
XX
KW   Embryonic flower gene-1; EMF-1; reproductive development; flowering;
KW   early flowering; uniform flowering; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   promoter        1. .3200
FT                   /*tag=  a
FT   CDS             4239. .8029
FT                   /*tag=  b
FT                   /note= "contains introns"
FT   exon            4239. .4333
FT                   /*tag=  c
FT                   /number= 1
FT   intron          4334. .4445
FT                   /*tag=  d
FT                   /number= 1
FT   exon            4446. .4621
FT                   /*tag=  e
FT                   /number= 2
FT   intron          4622. .4702
FT                   /*tag=  f
FT                   /number= 2
FT   exon            4703. .4821
FT                   /*tag=  g
FT                   /number= 3
FT   intron          4822. .4900
FT                   /*tag=  h
FT                   /number= 3
FT   exon            4901. .4954
FT                   /*tag=  i
FT                   /number= 4
FT   intron          4955. .5043
FT                   /*tag=  j
FT                   /number= 4
FT   exon            5044. .6305
FT                   /*tag=  k
FT                   /number= 5
FT   intron          6306. .6445
FT                   /*tag=  l
FT                   /number= 5
FT   exon            6446. .8029
FT                   /*tag=  m
FT                   /number= 6
XX
PN   WO200022132-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US023543.
XX
PR   09-OCT-1998;   98US-00169696.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Sung ZR,  Aubert D,  Chen L;
XX
DR   WPI; 2000-317987/27.
DR   P-PSDB; AAY84833.
XX
PT   An isolated nucleic acid molecule comprising embryonic flower genes is
PT   useful for modulating reproductive development in plants.
XX
PS   Disclosure; Page 34-38; 38pp; English.
XX
CC   The present sequence represents the Arabidopsis thaliana embryonic flower
CC   gene-1 (EMF-1). The EMF genes play an important role in reproductive
CC   development in plants. Control of the expression of the genes is
CC   therefore useful in controlling flowering and other functions in plants.
CC   The EMF-1 polynucleotides are useful for modulating reproductive
CC   development in plants. For example, enhanced expression of EMF-1 is
CC   useful to prevent flowering in grasses or vegetable crops. Controlling or
CC   inhibiting expression of the genes is useful to promote early flowering
CC   in forest trees and agronomical crops or to ensure uniform flowering in
CC   various agronomic crops or ornamentals
XX
SQ   Sequence 8646 BP; 2751 A; 1563 C; 1630 G; 2699 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              1.64e-11      Length:          8646
Score:                  298.00        Matches:         225
Percent Similarity:     35.9%         Conservative:    156
Best Local Similarity:  21.2%         Mismatches:      355
Query Match:            5.4%          Indels:          326
DB:                     3             Gaps:            51

US-09-828-068-2 (1-1057) x AAA14873 (1-8646)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
```


Db 5088 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5138
Qy 218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
Db 5139 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5198
Qy 235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------- 247
Db 5199 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5258
Qy 248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Db 5259 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5318
Qy 267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db 5319 GATAACGGGTCATCAGAAAGTATAAATCTTGCT---------- 5351
Qy 287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
Db 5352 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5405
Qy 304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db 5406 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA---------- 5447
Qy 324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db 5448 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5501
Qy 344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
Db 5502 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5561
Qy 361 AlaLysLeuCys---------- 364
Db 5562 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5621
Qy 365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db 5622 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5681
Qy 380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db 5682 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5735
Qy 400 CysGluAspAspArgSerThr----------IlePro 408
Db 5736 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5789
Qy 409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db 5790 TGTCCT----------CCGGGTACTCAGAGAACA------GAGAGGAAG 5822
Qy 429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db 5823 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5873
Qy 449 SerLeuMet----------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
Db 5874 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5933
Qy 463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db 5934 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5993
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
Db 5994 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6053
Qy 500 ----------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db 6054 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6113
Qy 516 ThrGlnArgCys----------SerSerLysGlyLysThrAlaGly---Leu 528
Db 6114 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6173
Qy 529 SerLysGly----------LysThr 533
Db 6174 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6233
Qy 534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db 6234 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6293
Qy 554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db 6294 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6353
Qy 573 r---------- 573
Db 6354 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6413
Qy 574 ----------Ala-LysValSerProAlaGluHisAspIleG 584
Db 6414 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6473
Qy 584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Db 6474 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6533
Qy 604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db 6534 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6590
Qy 624 euAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnA 644
Db 6591 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAA---GATGAAAGCAACA 6647
Qy 644 rgIleGlnSerLysThrThrAla------AspAspAspCysValIleValAlaAlaLysA 662
Db 6648 AACAGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATG 6707
Qy 662 spGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerL 680
Db 6708 AAACCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGT 6767
Qy 680 euAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluS 700
Db 6768 GTAGTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGA 6818
Qy 700 erProHis----------ProGlnAsn---------PheG 707
Db 6819 ACTCTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTC 6878
Qy 707 lnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaS 727
Db 6879 CTCCTACCCAAGAAAACCGA----------GCCA 6902
Qy 727 erSerProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisT 747
Db 6903 GCTCCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG---- 6958
Qy 747 rpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnS 767
Db 6959 --GGTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT---------- 7003
Qy 767 erProAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValA 784
Db 7004 -----TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------- 7039
Qy 784 spLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValI 804
Db 7040 -----GAAGCTTCTCATCCAATTTGGCCATCTTCC----------ATGA 7073

```
Qy       804 leAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheP 824
Db      7074 TACCACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---C 7130

Qy       824 roSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnV 844
Db      7131 CGGGTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTG 7184

Qy       844 alValLeuTyrProLysGlu---------------------------------------- 850
Db      7185 TTGCTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAAC 7244

Qy       851 ------------------------------------------------------------S 851
Db      7245 ATGCTGCTGGGGTTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCT 7304

Qy       851 erMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro- 870
Db      7305 CTATACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCG 7364

Qy       871 -----AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnS 887
Db      7365 CTGACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCA 7424

Qy       887 erGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuA 907
Db      7425 AAGAGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG---- 7480

Qy       907 snGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisA 927
Db      7481 -----CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCC 7529

Qy       927 rgProLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluI 946
Db      7530 GGAAGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAA 7583

Qy       946 leAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrG 966
Db      7584 ATGCTTCATGGAGTCCTCAT---------------------------------------- 7603

Qy       966 lyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetP 986
Db      7604 -----------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------T 7637

Qy       986 heSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp------- 1003
Db      7638 TTGCTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGA 7697

Qy      1004 --------PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyL 1021
Db      7698 AGTTCCAGCTGCTGGGAGCATCGAATTCC---------------------------- 7726

Qy      1021 ysMetValHisProLeu--------------AspArgPheValArgGlnAsp------- 1033
Db      7727 --ATGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAG 7784

Qy      1034 ------------------------------------------------------------I 1034
Db      7785 CAGAGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTG 7844

Qy      1034 leCysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrM 1054
Db      7845 TGTGCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACA 7904

Qy      1054 et 1054
Db      7905 TG 7906

RESULT 8
AAL43177
ID   AAL43177 standard; DNA; 8648 BP.
XX
AC   AAL43177;
XX
DT   16-AUG-2002  (first entry)
XX
DE   Arabidopsis thaliana embryonic flower 1 (EMF1) genomic sequence.
XX
KW   Embryonic flower 1; EMF1; gene; ds; early reproductive state transition;
KW   transgenic plant; flowering control; reproductive trait control.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   promoter        1. .3201
FT                   /*tag= a
FT   exon            3202. .3265
FT                   /*tag= c
FT                   /number= 1
FT   intron          3266. .4159
FT                   /*tag= d
FT                   /number= 1
FT   exon            4160. .4335
FT                   /*tag= e
FT                   /number= 2
FT   CDS             4241. .8031
FT                   /*tag= b
FT                   /product= "Arabidopsis thaliana EMF1 protein"
FT   intron          4336. .4447
FT                   /*tag= f
FT                   /number= 2
FT   exon            4448. .4623
FT                   /*tag= g
FT                   /number= 3
FT   intron          4624. .4703
FT                   /*tag= h
FT                   /number= 3
FT   exon            4704. .4823
FT                   /*tag= i
FT                   /number= 4
FT   intron          4824. .4902
FT                   /*tag= j
FT                   /number= 4
FT   exon            4903. .4956
FT                   /*tag= k
FT                   /number= 5
FT   intron          4957. .5045
FT                   /*tag= l
FT                   /number= 5
FT   exon            5046. .6307
FT                   /*tag= m
FT                   /number= 6
FT   intron          6308. .6447
FT                   /*tag= n
FT                   /number= 6
FT   exon            6448. .8065
FT                   /*tag= o
FT                   /number= 7
FT   intron          8066. .8300
FT                   /*tag= p
FT                   /number= 7
FT   exon            8301. .8648
FT                   /*tag= q
FT                   /number= 8
XX
PN   US6376751-B1.
XX
PD   23-APR-2002.
XX
PF   08-OCT-1999;   99US-00415946.
XX
PR   09-OCT-1998;   98US-00169696.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
```

PI Sung ZR, Aubert D, Chen L;
XX
DR WPI; 2002-453154/48.
DR P-PSDB; AAO15030.
XX
PT A polynucleotide which hybridizes to an embryonic flower 1 protein and
PT promotes early transition from vegetative to reproductive state when in a
PT plant is useful to control flowering in transgenic plants.
XX
PS Claim 3; Col 35-48; 30pp; English.
XX
CC The invention comprises the amino acid and coding sequence of the
CC Arabidopsis thaliana embryonic flower 1 (EMF1) protein. The EMF1
CC nucleotide of the invention promotes early transition from a vegetative
CC to a reproductive state when operably linked to a plant promoter and
CC introduced into a plant. The EMF1 nucleotide promotes early transition to
CC a reproductive state by hybridising to the EMF1 gene that is naturally
CC present within the plant. The EMF1 DNA sequence is useful for controlling
CC flowering and other reproductive traits in plants. The present DNA
CC sequence represents the Arabidopsis thaliana embryonic flower 1 (EMF1)
CC genomic sequence
XX
SQ Sequence 8648 BP; 2748 A; 1563 C; 1631 G; 2706 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.64e-11	Length:	8648
Score:	298.00	Matches:	225
Percent Similarity:	35.5%	Conservative:	152
Best Local Similarity:	21.2%	Mismatches:	360
Query Match:	5.4%	Indels:	324
DB:	6	Gaps:	50

US-09-828-068-2 (1-1057) x AAL43177 (1-8648)

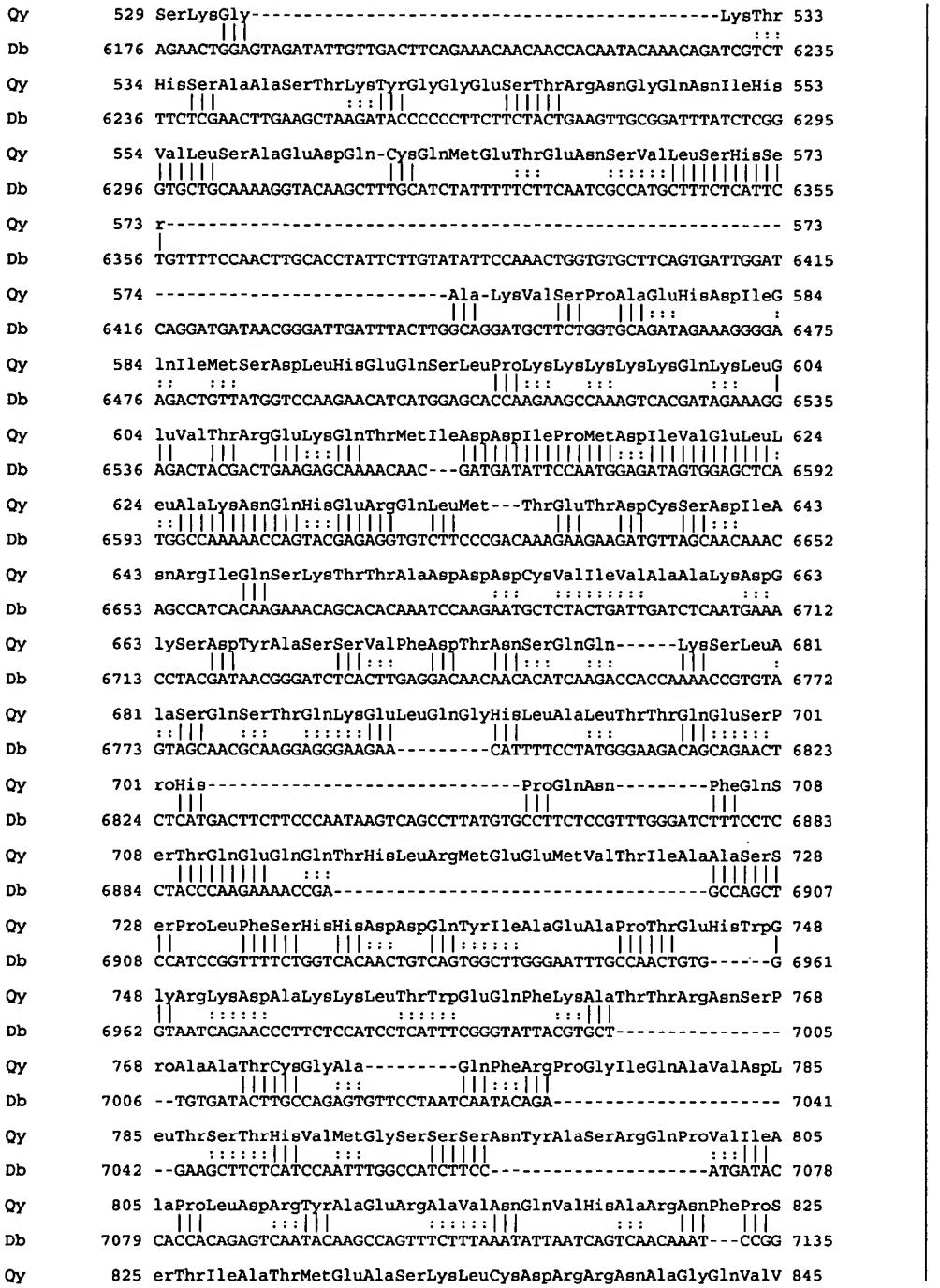
```
Qy       198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db      5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy       218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
Db      5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy       235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
Db      5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy       248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Db      5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy       267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db      5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT--------------------------- 5353

Qy       287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
Db      5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy       304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db      5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy       324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db      5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy       344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
Db      5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy       361 AlaLysLeuCys---------------------------------------------- 364
Db      5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy       365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db      5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy       380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db      5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy       400 CysGluAspAspArgSerThr--------------------------------IlePro 408
Db      5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy       409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db      5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy       429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db      5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy       449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
Db      5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy       463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db      5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy       483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
Db      5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy       500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db      6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy       516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
Db      6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy       529 SerLysGly-------------------------------------------------LysThr 533
Db      6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy       534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db      6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy       554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db      6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy       573 r------------------------------------------------------------ 573
Db      6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy       574 --------------------------Ala-LysValSerProAlaGluHisAspIleG 584
Db      6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475

Qy       584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
Db      6476 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6535

Qy       604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db      6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592

Qy       624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db      6593 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 6652
```

```
Qy       643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db      6653 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 6712

Qy       663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
Db      6713 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 6772

Qy       681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db      6773 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 6823

Qy       701 roHis----------------------------ProGlnAsn---------PheGlnS 708
Db      6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 6883

Qy       708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db      6884 CTACCCAAGAAAACCGA---------------------------------GCCAGCT 6907

Qy       728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db      6908 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 6961

Qy       748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db      6962 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT--------------- 7005

Qy       768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
Db      7006 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------------------- 7041

Qy       785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Db      7042 --GAAGCTTCTCATCCAATTTGGCCATCTTCC--------------------ATGATAC 7078

Qy       805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db      7079 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135

Qy       825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db      7136 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 7189

Qy       845 alLeuTyrProLysGlu---------------------------------------- 850
Db      7190 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 7249

Qy       851 ----------------------------------------------------SerM 852
Db      7250 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309

Qy       852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
Db      7310 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 7369

Qy       871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db      7370 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 7429

Qy       888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
Db      7430 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 7482

Qy       908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db      7483 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 7534

Qy       928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db      7535 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 7588

Qy       947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db      7589 CTTCATGGAGTCCTCAT---------------------------------------- 7605

Qy       967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db      7606 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------TTTG 7642

Qy       987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
Db      7643 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 7702

Qy      1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db      7703 TCCAGCTGCTGGGAGCATCGAATTCC--------------------------------A 7729

Qy      1022 etValHisProLeu--------------AspArgPheValArgGlnAsp---------- 1033
Db      7730 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 7789

Qy      1034 ------------------------------------------------------IleC 1035
Db      7790 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 7849

Qy      1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db      7850 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 7908
```

RESULT 9
ABQ77909
ID ABQ77909 standard; DNA; 8648 BP.
XX
AC ABQ77909;
XX
DT 24-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana EMF1 gene.
XX
KW EMF1; embryonic flower 1; rice OsEMF1 homologue; plant; floral repressor;
KW reproductive development; flower development; transgenic plant;
KW antisense suppression; transgenic; reciprocal negative interaction;
KW flower meristem identity gene; flowering time; shoot development;
KW seed yield; agriculture; gene; ds.
XX
OS Arabidopsis thaliana.
XX

FH	Key	Location/Qualifiers
FT	CDS	4241. .8031
FT		/*tag= a
FT		/product= "EMF1 protein"
FT	exon	4241. .4335
FT		/*tag= b
FT		/number= 1
FT	intron	4336. .4447
FT		/*tag= c
FT		/number= 1
FT	exon	4448. .4623
FT		/*tag= d
FT		/number= 2
FT	intron	4624. .4703
FT		/*tag= e
FT		/number= 2
FT	exon	4704. .4823
FT		/*tag= f
FT		/number= 3
FT	intron	4824. .4902
FT		/*tag= g
FT		/number= 3
FT	exon	4903. .4956
FT		/*tag= h
FT		/number= 4
FT	intron	4957. .5045
FT		/*tag= i
FT		/number= 4

```
FT   exon            5046. .6307
FT                   /*tag= j
FT                   /number= 5
FT   intron          6308. .6447
FT                   /*tag= k
FT                   /number= 5
FT   exon            6448. .8031
FT                   /*tag= l
FT                   /number= 6
XX
PN   WO200280659-A1.
XX
PD   17-OCT-2002.
XX
PF   05-APR-2002; 2002WO-US012675.
XX
PR   06-APR-2001; 2001US-00828068.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Moon Y, Chen L, Sung ZR;
XX
DR   WPI; 2003-046831/04.
DR   P-PSDB; ABB99878.
XX
PT   New isolated OsEMF1 nucleic acid molecule and encoded polypeptide, useful
PT   for plant genetic engineering, in particular controlling reproductive
PT   development in rice.
XX
PS   Example 2; Page; 47pp; English.
XX
CC   The invention relates to a rice EMF (embryonic flower) gene designated
CC   OsEMF1 (cDNA given in ABQ77908) and its encoded protein (ABB99878). The
CC   OsEMF1 protein has a molecular weight of 116.4 kD and exhibits 37%
CC   homology and 20% identity with Arabidopsis thaliana EMF1 protein
CC   (ABB99879). OsEMF1, like other EMF gene products, acts as a floral
CC   repressor, suppressing the transition from vegetative growth to
CC   reproductive development. It also delays the inflorescence to flower
CC   transition, indicating that there is a reciprocal negative interaction
CC   between OsEMF and flower meristem identity genes. The invention also
CC   encompasses a transgenic plant comprising an expression cassette
CC   containing an OsEMF1 nucleic acid (particularly in an antisense
CC   orientation) under the control of a plant promoter. OsEMF1 nucleic acids
CC   and proteins may be used for controlling reproductive development in
CC   plants, particularly monocotyledonous plants and especially rice. In
CC   particular, OsEMF1 nucleic acids and proteins may be used to control
CC   flowering time, shoot development and seed yield. Controlling or
CC   inhibiting the expression of genes which mediate these processes enables
CC   new varieties of rice with different flowering times and seed yield to be
CC   developed. The present sequence represents the Arabidopsis thaliana EMF1
CC   gene which is referred to in an exemplification of the invention. Note:
CC   The present sequence is not given in the specification, but was obtained
CC   from GenBank (accession number AF319968)
XX
SQ   Sequence 8648 BP; 2748 A; 1563 C; 1631 G; 2706 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.64e-11      Length:          8648
Score:                    298.00        Matches:         225
Percent Similarity:       35.5%         Conservative:    152
Best Local Similarity:    21.2%         Mismatches:      360
Query Match:              5.4%          Indels:          324
DB:                       8             Gaps:            50

US-09-828-068-2 (1-1057) x ABQ77909 (1-8648)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
              ::::|||     |||      :::   |||     |||   |||        |||
Db       5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
              |||   |||::::::             |||:::|||:::   :::        |||
Db       5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
                    :::         :::         |||:::
Db       5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
                 |||   ||||||          |||      :::   :::             |||
Db       5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy        267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
              :::||||||   :::|||         :::|||
Db       5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT---------------------------- 5353

Qy        287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
                  |||||||||   :::|||           :::|||                  |||
Db       5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy        304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
                  :::   |||         |||             :::
Db       5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA---------------- 5449

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
              ||||||      ||||||:::   |||          ||||||      :::|||||:::
Db       5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy        344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
                 :::      :::   |||:::      :::              |||:::
Db       5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy        361 AlaLysLeuCys---------------------------------------------- 364
              :::|||   |||
Db       5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy        365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
                           :::::::::   ::::::   :::::::::   |||   :::
Db       5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy        380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
              :::         |||   |||:::      |||      :::|||:::   |||||||||
Db       5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy        400 CysGluAspAspArgSerThr--------------------------------IlePro 408
                :::      |||||||||                                 :::|||
Db       5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy        409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
                  |||         |||   :::      |||      |||
Db       5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy        429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
              |||      |||   |||||||||:::            |||:::|||:::|||
Db       5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy        449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
              |||   :::            |||   ::::::|||   :::            ||||||
Db       5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
                  :::|||||:::                 |||:::   |||:::|||
Db       5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
                    :::         :::                      |||
Db       5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy        500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
                         |||::::::|||      |||                |||    :::
Db       6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy        516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
                    |||                  |||||||||      :::   |||   |||
Db       6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175
```



```
RESULT 10
AAA14872
ID   AAA14872 standard; DNA; 17341 BP.
XX
AC   AAA14872;
XX
DT   08-AUG-2000  (first entry)
XX
DE   Genomic DNA sequence from Arabidopsis thaliana CD82 clone.
XX
KW   Embryonic flower gene-1; EMF-1; reproductive development; flowering;
KW   early flowering; uniform flowering; ss.
XX
OS   Arabidopsis thaliana.
```

XX
PN WO200022132-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023543.
XX
PR 09-OCT-1998; 98US-00169696.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Sung ZR, Aubert D, Chen L;
XX
DR WPI; 2000-317987/27.
XX
PT An isolated nucleic acid molecule comprising embryonic flower genes is
PT useful for modulating reproductive development in plants.
XX
PS Claim 3; Page 30-33; 38pp; English.
XX
CC The present sequence represents the genomic sequence of the Arabidopsis
CC thaliana genomic DNA from CD82 clone. The sequence comprises the
CC embryonic flower gene-1 (EMF-1). The EMF genes play an important role in
CC reproductive development in plants. Control of the expression of the
CC genes is therefore useful in controlling flowering and other functions in
CC plants. The EMF-1 polynucleotides are useful for modulating reproductive
CC development in plants. For example, enhanced expression of EMF-1 is
CC useful to prevent flowering in grasses or vegetable crops. Controlling or
CC inhibiting expression of the genes is useful to promote early flowering
CC in forest trees and agronomical crops or to ensure uniform flowering in
CC various agronomic crops or ornamentals
XX
SQ Sequence 17341 BP; 5468 A; 3114 C; 3409 G; 5350 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.12e-11	Length:	17341
Score:	298.00	Matches:	225
Percent Similarity:	35.5%	Conservative:	152
Best Local Similarity:	21.2%	Mismatches:	360
Query Match:	5.4%	Indels:	324
DB:	3	Gaps:	50

US-09-828-068-2 (1-1057) x AAA14872 (1-17341)

```
Qy       198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
             :::|||      |||      :::   |||      |||  |||         |||
Db      5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy       218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
             |||   |||::::::               |||:::|||:::  :::      |||
Db      5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy       235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla--------------------- 247
                 :::     :::       |||:::
Db      5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy       248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
                |||   ||||||          |||      :::   :::          |||
Db      5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy       267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
             :::||||||   :::|||         :::|||
Db      5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT--------------------------- 5353

Qy       287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
                 ||||||||  :::|||          :::|||                  |||
Db      5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy       304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
                 :::   |||       |||          :::
Db      5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy       324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
```

```
             ||||||      ||||||:::  |||        ||||||      :::||||||:::
Db      5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy       344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
                :::      :::   |||:::      :::                |||:::
Db      5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy       361 AlaLysLeuCys---------------------------------------------- 364
             :::|||   |||
Db      5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy       365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
                          :::::::::   ::::::   :::::::::    |||   :::
Db      5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy       380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
             :::          |||   |||:::      |||       :::|||:::   |||||||||
Db      5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy       400 CysGluAspAspArgSerThr--------------------------------IlePro 408
                :::      ||||||||                                  :::|||
Db      5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy       409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
                |||                  |||  :::     |||    |||
Db      5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy       429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
             |||          |||   |||||||||:::            |||:::|||:::|||
Db      5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy       449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
             |||   :::            |||   :::::::|||   :::           ||||||
Db      5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy       463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
                 :::||||||:::                |||:::   |||:::|||
Db      5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy       483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
                      :::      :::                      |||
Db      5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy       500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
                        |||::::::|||      |||                |||      :::
Db      6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy       516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
                 |||                     |||||||||     :::   |||   |||
Db      6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy       529 SerLysGly---------------------------------------------LysThr 533
                 |||                                                :::
Db      6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy       534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
                |||            ::|||        ||||||
Db      6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy       554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
             ||||||              |||      :::      ::::::||||||||||
Db      6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy       573 r------------------------------------------------------------ 573
             |
Db      6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy       574 --------------------------Ala-LysValSerProAlaGluHisAspIleG 584
                                        |||      |||   ||:::      :
Db      6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475

Qy       584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
             ::   :::                        |||:::   :::         :::    |
```

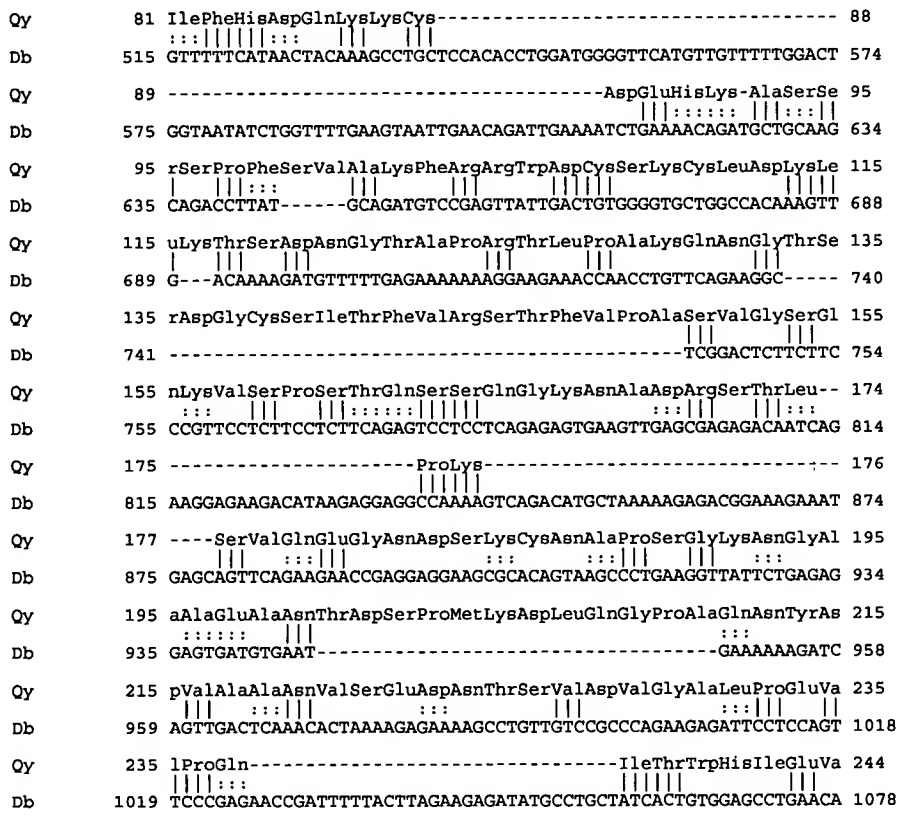
```
Db       6476 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6535

Qy        604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
              ||   |||   |||:::|||          |||||||||||||:::||||||||||||:
Db       6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592

Qy        624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
              ::||||||||||||:::|||||   |||        |||   |||   |||:::
Db       6593 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 6652

Qy        643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
                 |||                    :::   :::::::::          :::
Db       6653 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 6712

Qy        663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
                 |||          |||:::   |||   |||:::   :::      |||      :
Db       6713 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 6772

Qy        681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
              ::|||   :::   ::::::|||          |||      :::      |||::::::
Db       6773 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 6823

Qy        701 roHis----------------------------ProGlnAsn---------PheGlnS 708
                 |||                          |||               |||
Db       6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 6883

Qy        708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
                 |||||||||   :::                                  |||||||
Db       6884 CTACCCAAGAAAACCGA--------------------------------GCCAGCT 6907

Qy        728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
              ||      ||||||   |||:::   |||::::::          ||||||         |
Db       6908 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 6961

Qy        748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
              ||   ::::::              ::::::          :::|||
Db       6962 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT--------------- 7005

Qy        768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
                 ||||||   :::          |||:::|||
Db       7006 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------------------- 7041

Qy        785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
                 ::::::|||   :::      ||||||                  :::|||
Db       7042 --GAAGCTTCTCATCCAATTTGGCCATCTTCC-------------------ATGATAC 7078

Qy        805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
                 |||      :::|||          :::::::|||          :::   |||   |||
Db       7079 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135

Qy        825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
                 |||::::::      :::||||||          :::      |||          |||
Db       7136 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 7189

Qy        845 alLeuTyrProLysGlu------------------------------------------ 850
                       |||:::
Db       7190 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 7249

Qy        851 ----------------------------------------------------------SerM 852
                                                                       |||:
Db       7250 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309

Qy        852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
              ::||||||   |||||||||   ::::::||||||          ::::::   |||
Db       7310 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 7369

Qy        871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
                 :::|||   :::      ::::::                        :::   :
Db       7370 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 7429

Qy        888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
              :::::          |||          |||:::   ::::::|||   :::
Db       7430 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 7482

Qy        908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
                 :::||||||       |||:::      |||:::             :::      |||
Db       7483 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 7534

Qy        928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
                       |||      |||   :::|||          |||   |||   :::    |
Db       7535 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 7588

Qy        947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
              ||:::||||||   :::
Db       7589 CTTCATGGAGTCCTCAT---------------------------------------------- 7605

Qy        967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
                       ||||||      :::   :::      :::                  |||:
Db       7606 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------TTTG 7642

Qy        987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
              ::      :::|||      :::          :::||||||   |||
Db       7643 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 7702

Qy       1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
                   |||   |||   ||||||                                    |
Db       7703 TCCAGCTGCTGGGAGCATCGAATTCC--------------------------------A 7729

Qy       1022 etValHisProLeu---------------AspArgPheValArgGlnAsp---------- 1033
              ||:::   ||||||               |||:::      :::|||
Db       7730 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 7789

Qy       1034 ----------------------------------------------------------IleC 1035
                                                                       :::|
Db       7790 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 7849

Qy       1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
              ||      |||:::||||||||||||||||||          |||   ||||||
Db       7850 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 7908
```

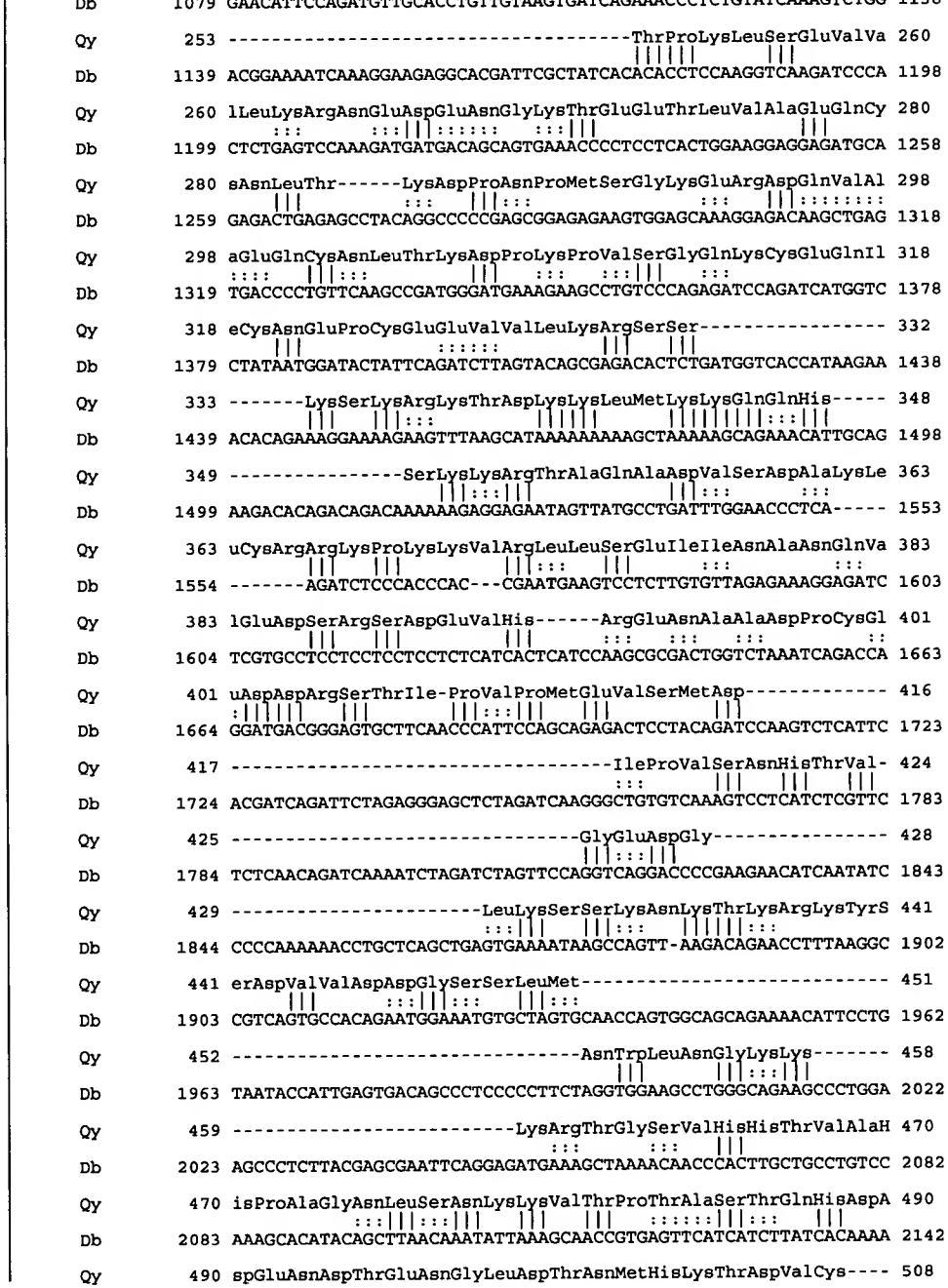
RESULT 11
AAL43176
ID AAL43176 standard; DNA; 17341 BP.
XX
AC AAL43176;
XX
DT 16-AUG-2002 (first entry)
XX
DE Arabidopsis thaliana embryonic flower 1 gene contained within a vector.
XX
KW Embryonic flower 1; EMF1; gene; ds; early reproductive state transition;
KW transgenic plant; flowering control; reproductive trait control;
KW pBluescript vector.
XX
OS Arabidopsis thaliana.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT promoter 1. .3201
FT /*tag= a
FT CDS 4241. .8031
FT /*tag= b
FT /product= "Arabidopsis EMF1 protein"
FT /note= "The CDS contains introns"
XX
PN US6376751-B1.
XX
PD 23-APR-2002.
XX
PF 08-OCT-1999; 99US-00415946.
XX
PR 09-OCT-1998; 98US-00169696.
XX
PA (REGC) UNIV CALIFORNIA.
XX

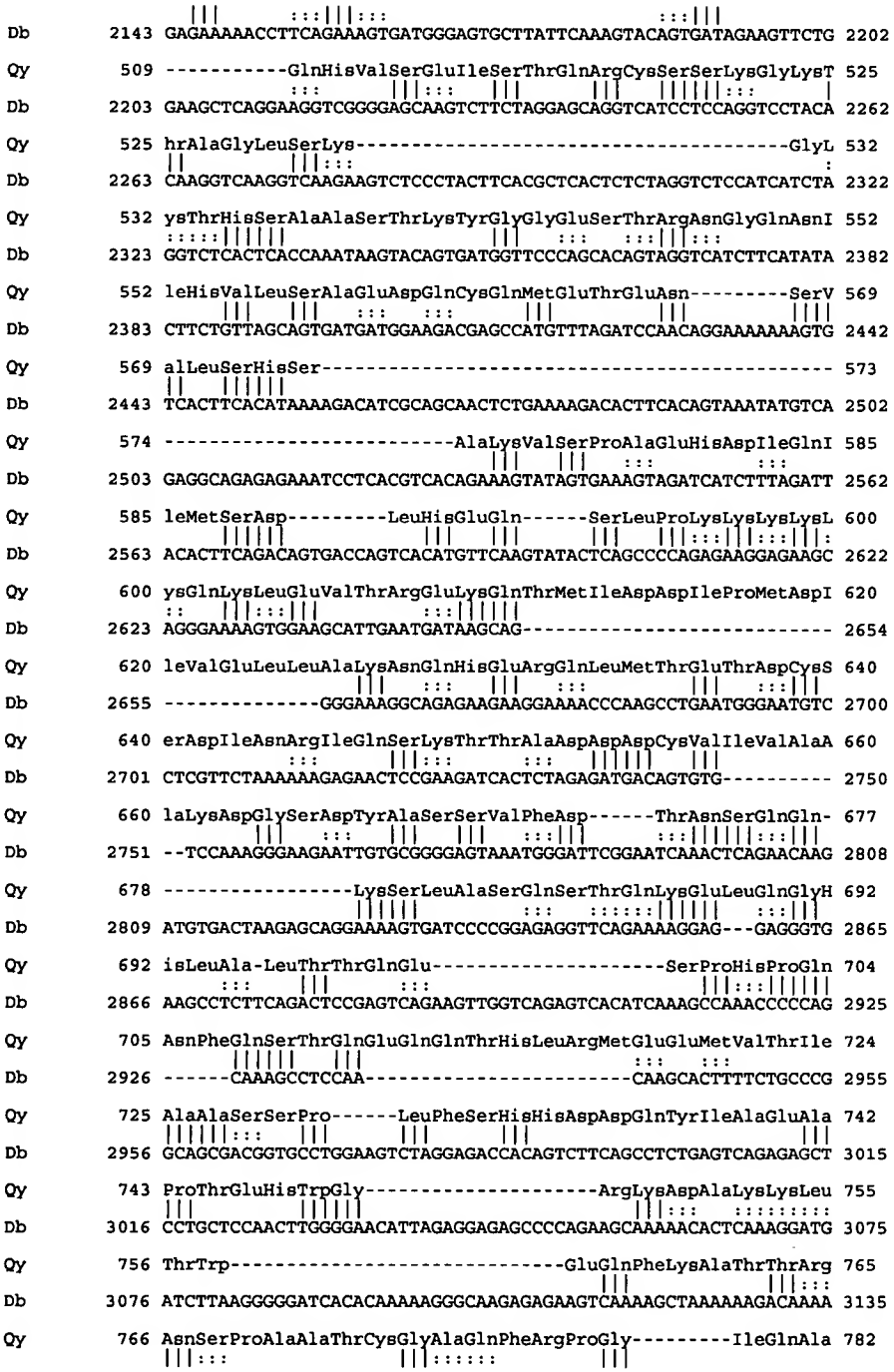
```
PI   Sung ZR,  Aubert D,  Chen L;
XX
DR   WPI; 2002-453154/48.
DR   P-PSDB; AAO15030.
XX
PT   A polynucleotide which hybridizes to an embryonic flower 1 protein and
PT   promotes early transition from vegetative to reproductive state when in a
PT   plant is useful to control flowering in transgenic plants.
XX
PS   Example 1; Col 19-34; 30pp; English.
XX
CC   The invention comprises the amino acid and coding sequence of the
CC   Arabidopsis thaliana embryonic flower 1 (EMF1) protein. The EMF1
CC   nucleotide of the invention promotes early transition from a vegetative
CC   to a reproductive state when operably linked to a plant promoter and
CC   introduced into a plant. The EMF1 nucleotide promotes early transition to
CC   a reproductive state by hybridising to the EMF1 gene that is naturally
CC   present within the plant. The EMF1 DNA sequence is useful for controlling
CC   flowering and other reproductive traits in plants. The present DNA
CC   sequence represents the Arabidopsis thaliana embryonic flower 1 (EMF1)
CC   gene sequence contained within a pBluescript vector
XX
SQ   Sequence 17341 BP; 5468 A; 3114 C; 3409 G; 5350 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.12e-11        Length:         17341
Score:                  298.00          Matches:        225
Percent Similarity:     35.5%           Conservative:   152
Best Local Similarity:  21.2%           Mismatches:     360
Query Match:            5.4%            Indels:         324
DB:                     6               Gaps:           50

US-09-828-068-2 (1-1057) x AAL43176 (1-17341)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
              :::|||     |||      :::   |||      |||   |||       |||
Db       5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
                |||   |||::::::           |||:::|||:::   :::      |||
Db       5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
                       :::        :::         |||:::
Db       5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
                 |||   ||||||         |||      :::   :::            |||
Db       5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy        267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
               :::||||||   :::|||         :::|||
Db       5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT------------------------- 5353

Qy        287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
                  |||||||||   :::|||           :::|||                |||
Db       5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy        304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
                  :::   |||      |||          :::
Db       5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
                  ||||||      ||||||:::   |||          ||||||      :::||||||:::
Db       5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy        344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
                  :::      :::   |||:::      :::               |||:::
Db       5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy        361 AlaLysLeuCys------------------------------------------------ 364
                  :::|||   |||
Db       5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy        365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
                           :::::::::   ::::::   :::::::::    |||    :::
Db       5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy        380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
                  :::      |||   |||:::      |||      :::|||:::   |||||||||
Db       5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy        400 CysGluAspAspArgSerThr--------------------------------IlePro 408
                  :::      |||||||||                                :::|||
Db       5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy        409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
                  |||                  |||   :::      |||      |||
Db       5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy        429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
                  |||         |||   |||||||||:::           |||:::|||:::|||
Db       5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy        449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
                  |||   :::            |||   ::::::|||   :::            ||||||
Db       5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
                     :::||||||:::                |||:::   |||:::|||
Db       5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
                        :::       :::                |||
Db       5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy        500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
                           |||::::::|||       |||                |||      :::
Db       6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy        516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
                  |||                     |||||||||       :::   |||   |||
Db       6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy        529 SerLysGly-------------------------------------------LysThr 533
                  |||                                                :::
Db       6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy        534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
                  |||      :::|||      ||||||
Db       6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy        554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
                  ||||||                |||         :::      ::::::||||||||||
Db       6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy        573 r------------------------------------------------------------ 573
                  |
Db       6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy        574 --------------------------Ala-LysValSerProAlaGluHisAspIleG 584
                                       |||      |||   |||:::          :
Db       6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475

Qy        584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
                  ::   :::                        |||:::   :::          :::   |
Db       6476 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6535

Qy        604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
                  ||   |||   |||:::|||       ||||||||||||||:::||||||||||||:
Db       6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592

Qy        624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
                  ::||||||||||:::||||||   |||          |||   |||   |||:::
Db       6593 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 6652
```

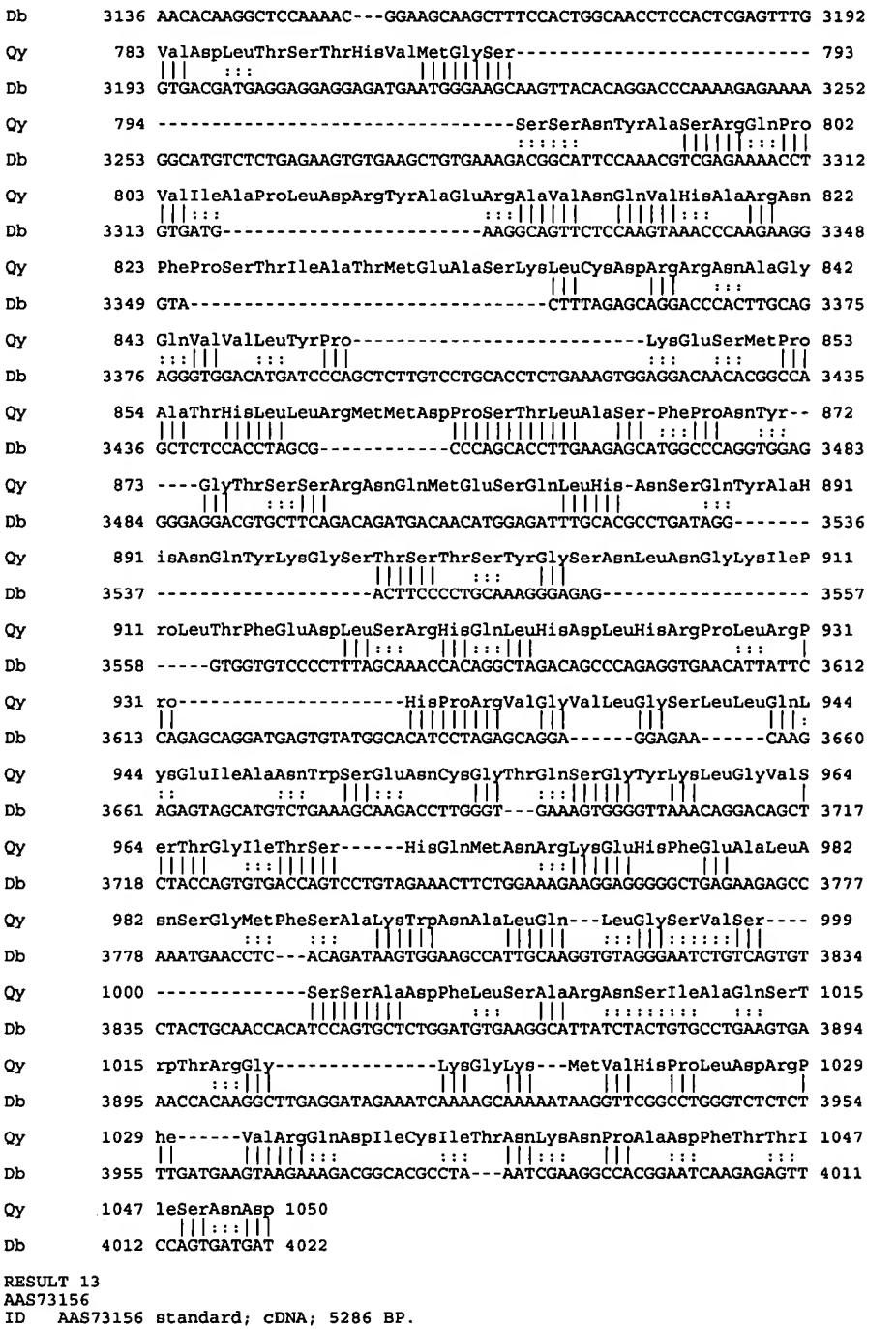
```
Qy       643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db      6653 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 6712

Qy       663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
Db      6713 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 6772

Qy       681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db      6773 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 6823

Qy       701 roHis----------------------------ProGlnAsn---------PheGlnS 708
Db      6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 6883

Qy       708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db      6884 CTACCCAAGAAAACCGA---------------------------------GCCAGCT 6907

Qy       728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db      6908 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 6961

Qy       748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db      6962 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT---------------- 7005

Qy       768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
Db      7006 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------------------- 7041

Qy       785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Db      7042 --GAAGCTTCTCATCCAATTTGGCCATCTTCC--------------------ATGATAC 7078

Qy       805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db      7079 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135

Qy       825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db      7136 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 7189

Qy       845 alLeuTyrProLysGlu---------------------------------------- 850
Db      7190 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 7249

Qy       851 ----------------------------------------------------SerM 852
Db      7250 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309

Qy       852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
Db      7310 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 7369

Qy       871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db      7370 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 7429

Qy       888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
Db      7430 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 7482

Qy       908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db      7483 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 7534

Qy       928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db      7535 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 7588

Qy       947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db      7589 CTTCATGGAGTCCTCAT---------------------------------------- 7605

Qy       967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db      7606 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------TTTG 7642

Qy       987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
Db      7643 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 7702

Qy      1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db      7703 TCCAGCTGCTGGGAGCATCGAATTCC--------------------------------A 7729

Qy      1022 etValHisProLeu--------------AspArgPheValArgGlnAsp---------- 1033
Db      7730 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 7789

Qy      1034 ------------------------------------------------------IleC 1035
Db      7790 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 7849

Qy      1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db      7850 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 7908

RESULT 12
ADZ61911
ID   ADZ61911 standard; DNA; 7036 BP.
XX
AC   ADZ61911;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Murine Nktr gene, SEQ ID 1635.
XX
KW   Drug screening; Peroxisome Proliferator-Activated Receptor gamma;
KW   PPAR gamma; gene; ds.
XX
OS   Mus musculus.
XX
PN   US2005084872-A1.
XX
PD   21-APR-2005.
XX
PF   23-JAN-2004; 2004US-00764420.
XX
PR   24-JAN-2003; 2003US-0442797P.
PR   30-MAY-2003; 2003US-0474413P.
XX
PA   (LUMP/) LUM P Y.
PA   (TANY/) TAN Y.
PA   (DAIH/) DAI H.
PA   (MUIS/) MUISE E S.
PA   (BERG/) BERGER J P.
PA   (THOM/) THOMPSON J R.
XX
PI   Lum PY,  Tan Y,  Dai H,  Muise ES,  Berger JP,  Thompson JR;
XX
DR   WPI; 2005-313963/32.
DR   REFSEQ; NM_010918.
XX
PT   Determining whether agent has defined biological activity, by making
PT   comparison e.g. comparing efficacy value, toxicity value and classifier
PT   value of agent to reference values, and using results to determine
PT   whether agent has activity.
XX
PS   Example 3; SEQ ID NO 1635; 51pp; English.
XX
CC   The present invention relates to a methods for screening biologically
CC   active agents, such as candidate drugs, to determine whether the agent
CC   possesses a defined biological activity. The methods involve making one
```

```
CC   or more comparisons chosen from comparing efficacy value of agent to
CC   reference efficacy value, comparing a toxicity value of the agent to
CC   reference toxicity value, comparing a classifier value of the agent to
CC   reference classifier value, and using the comparison result(s) to
CC   determine whether the agent possesses the defined biological activity.
CC   Also claimed is a population of oligonucleotide probes (I) specific for
CC   measuring the expression levels of members of a classifier population of
CC   genes or a toxicity-related population of genes. (I) are useful for
CC   measuring the expression levels of genes that are useful for identifying
CC   agonists or partial agonists of Peroxisome Proliferator-Activated
CC   Receptor (PPAR) gamma. ADZ62007-ADZ62272 are oligonucleotide probes which
CC   are useful in the method of the invention to measure the expression
CC   pattern of the classifier-related population of mouse genes ADZ61748-
CC   ADZ62006. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from USPTO at seqdata.uspto.gov/sequence.html?DocID=20050084872.
XX
SQ   Sequence 7036 BP; 2081 A; 1579 C; 1709 G; 1667 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.05e-07        Length:          7036
Score:                    241.50          Matches:         266
Percent Similarity:       33.3%           Conservative:    162
Best Local Similarity:    20.7%           Mismatches:      428
Query Match:              4.4%            Indels:          432
DB:                       14              Gaps:            60

US-09-828-068-2 (1-1057) x ADZ61911 (1-7036)
```



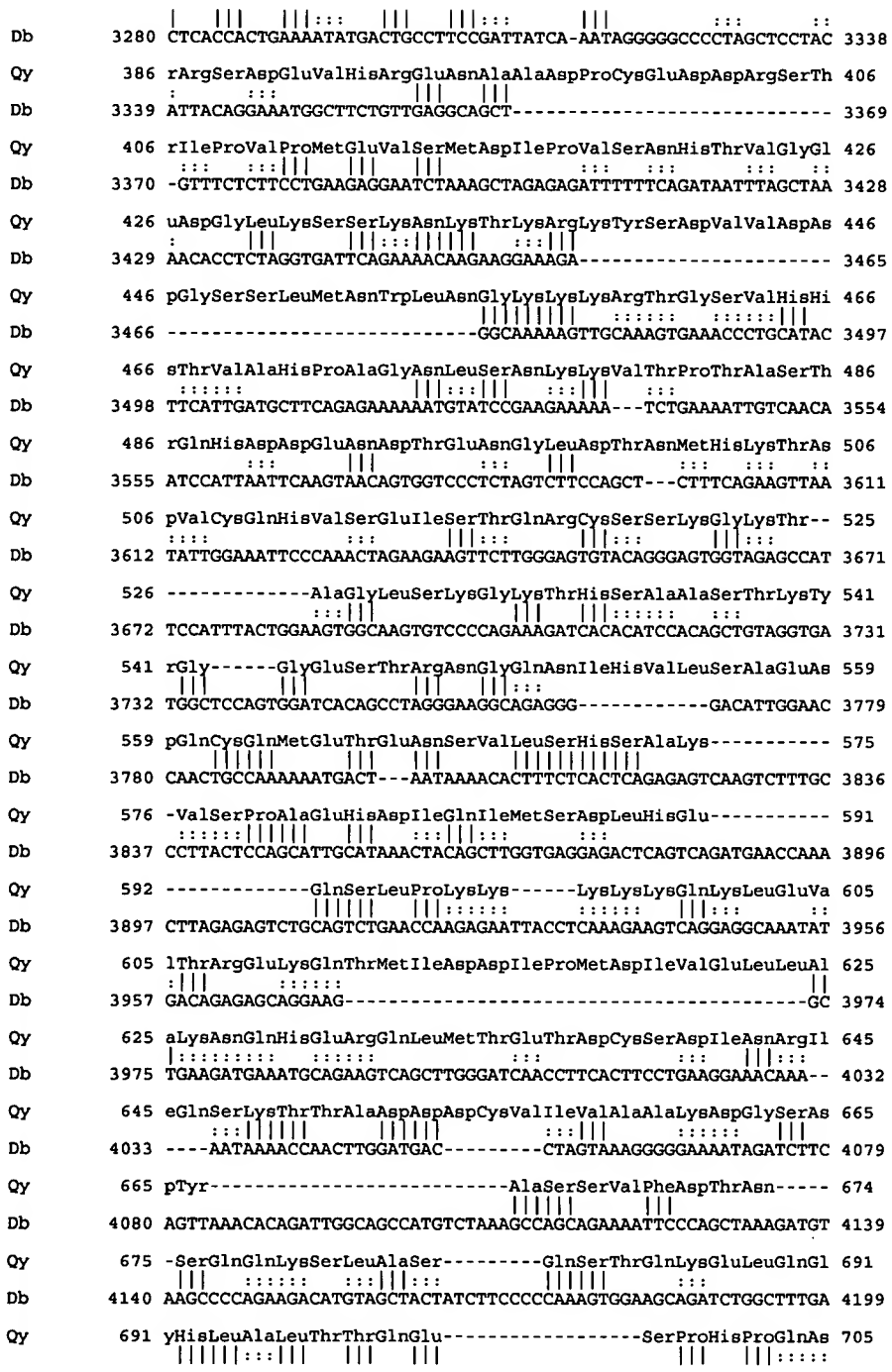






RESULT 13
AAS73156
ID AAS73156 standard; cDNA; 5286 BP.

```
XX
AC   AAS73156;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8960.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08969.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8960; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 5286 BP; 1786 A; 1178 C; 1073 G; 1249 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.1e-06        Length:          5286
Score:                    223.50         Matches:         210
Percent Similarity:       34.2%          Conservative:    173
Best Local Similarity:    18.8%          Mismatches:      412
Query Match:              4.0%           Indels:          325
DB:                       5              Gaps:            52

US-09-828-068-2 (1-1057) x AAS73156 (1-5286)

Qy          83 HisAspGln---------LysLysCysAspGluHisLysAlaSerSerSerProPheSer 99
               |||||:::          ||||||     |||          :::
Db        2236 CATGATGAAGTGGTTGATGTCAAATGCCATTCACACTCTCCTTTTAGGAATGAAAGAGGA 2295

Qy         100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
               |||   |||                 |||::::::||||||      ::::::
Db        2296 AAAGGAAAAATAAGGCATCAT------ATATCCTGTATTGAAAAGTTAAGCAAAACAGAA 2349

Qy         120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
               :::  :::   |||                 |||||||||      |||
Db        2350 AGTATATCAGTACCC-------------------------ACCAGTGATCACAGGAGC 2382

Qy         140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer--- 158
               :::                  :::   |||:::     :::   |||||||||
Db        2383 CTC-------------------ATTGAAGCAAATCAAAGCAATTCCAAAGTTTCTGAA 2421

Qy         159 ----------------------ProSerThrGlnSerSer---------GlnGlyLys 167
                                    |||         ||||||           |||:::
Db        2422 CTTGACACAATTTATTGTACCTTGCCAAGAAAATCAAGCAGTTTTCTCATACATGGCAGG 2481

Qy         168 AsnAlaAspArgSerThrLeuProLysSerValGlnGluGly------------------ 181
                  :::           :::     |||::::::   |||
Db        2482 CAGTCAGGAAGTAAAATAATGGCTGCTTCATTGAGGAATGGGCCACCTCCCTTCCAAATC 2541

Qy         182 -------------------------------------AsnAspSerLysCysAsnAla 188
                                                   |||          ::::::
Db        2542 AAAAATAATGTGGAAGATGCAATGGGGAACTATATGTTAAACAAATTTAGTCCCAGTTCT 2601

Qy         189 ProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln 208
               |||         |||     ::::::     :::||||||           |||:::
Db        2602 CCTGAGTCAGCGAATGAATGTTCCAAAGTCCTTTCAGACTCAGCC---------CTGGAA 2652

Qy         209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
               |||                 :::              |||      :::||||||||||
Db        2653 GCACCTGAAGCCACAGAGAGAATGACAAATGTAAAAAGCAGTGGATCTACTTCCGTTAGA 2712

Qy         229 ValGlyAlaLeuPro--------------------GluValProGlnIleThrTrpHis 241
                  |||   ||||||                        |||              |||
Db        2713 AAAGGACCACTTCCATTCCTCATCAACAGGGCTATGTCATGTCCCTCAGGGGAGCCACAT 2772

Qy         242 IleGlu-----------------------------ValAsnGlyAlaAspGlnProPro 251
                                                  :::|||   |||
Db        2773 GCCTCAACTGGAAGAGAAGGAAGAAAAAAGCCATTGACCTCAGGCATGGATGCTTCTGAG 2832

Qy         252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
                  ||||||:::      |||   ::::::          |||   ::::::
Db        2833 CTAACACCAAGGGCTTGGGAGAGAATCATTAGCCCTGTGGAAAGTGACTCATCT------ 2886

Qy         272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys---------------------- 284
                                   |||:::|||||||||
Db        2887 ---------------GTTAGAGATTGTTCTTTAACCAAAAGACAACACCAAAAGGAAAAC 2931

Qy         285 ---------------AspProAsnProMetSerGlyLysGluArgAspGlnValAlaGlu 299
                             :::         :::|||||||||||||||   :::
Db        2932 TTCCAAGAATACACTGAGAAAGAGGACTTGTCAGGAAAAGAACGTGGGAAAACATTACAT 2991

Qy         300 GlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCys 319
               :::          |||                |||||||||   :::              |||
Db        2992 AAAGTTAAG---ACGACTAGTACGTTTTCTGTTTCTGGTGATGAAGATAATGTAAAATGT 3048

Qy         320 AsnGluPro-CysGluGluValValLeuLysArgSerSerLysSerLys----------- 335
                  |||       :::   :::   |||   |||   :::   :::|||
Db        3049 CTGGAGGTGGTCTCAATATATTACACTCTACCGAGGAAACCCAGCAAAAAATTCTGTAAC 3108

Qy         335 ------------------------------------------------------------ 335

Db        3109 CTCCTTCAACAGTATACACAAAATACTAATTTACTTATAGAATCACCTCAAGTGGAGACT 3168

Qy         336 ----------------------ArgLysThrAspLysLysLeuMet------LysLysGl 346
                                    :::||||||::::::   ::::::            ||
Db        3169 GAAACATTTCCTAACGCTTTAGAAAAAGACAAACAGAATTATTCTACACGAGAGCAGTCA 3228

Qy         346 nGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgAr 366
               |:::|||       :::     :::                :::   |||   |||||
Db        3229 GGAACACCTTCATGTGAAAATCTAAAGATGT---------CCGTCAACTCTGATCAGACG 3279

Qy         366 gLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSe 386
```


Db 3280 CTCACCACTGAAAATATGACTGCCTTCCGATTATCA-AATAGGGGGCCCCTAGCTCCTAC 3338
Qy 386 rArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerTh 406
Db 3339 ATTACAGGAAATGGCTTCTGTTGAGGCAGCT---------- 3369
Qy 406 rIleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGl 426
Db 3370 -GTTTCTCTTCCTGAAGAGGAATCTAAAGCTAGAGAGATTTTTTCAGATAATTTAGCTAA 3428
Qy 426 uAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAs 446
Db 3429 AACACCTCTAGGTGATTCAGAAAACAAGAAGGAAAGA---------- 3465
Qy 446 pGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHi 466
Db 3466 ----------GGCAAAAAGTTGCAAAGTGAAACCCTGCATAC 3497
Qy 466 sThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerTh 486
Db 3498 TTCATTGATGCTTCAGAGAAAAAATGTATCCGAAGAAAAA---TCTGAAAATTGTCAACA 3554
Qy 486 rGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAs 506
Db 3555 ATCCATTAATTCAAGTAACAGTGGTCCCTCTAGTCTTCCAGCT---CTTTCAGAAGTTAA 3611
Qy 506 pValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThr-- 525
Db 3612 TATTGGAAATTCCCAAACTAGAAGAAGTTCTTGGGAGTGTACAGGGAGTGGTAGAGCCAT 3671
Qy 526 ----------AlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTy 541
Db 3672 TCCATTTACTGGAAGTGGCAAGTGTCCCCAGAAAGATCACACATCCACAGCTGTAGGTGA 3731
Qy 541 rGly------GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAs 559
Db 3732 TGGCTCCAGTGGATCACAGCCTAGGGAAGGCAGAGGG----------GACATTGGAAC 3779
Qy 559 pGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys---------- 575
Db 3780 CAACTGCCAAAAAATGACT---AATAAAACACTTTCTCACTCAGAGAGTCAAGTCTTTGC 3836
Qy 576 -ValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGlu---------- 591
Db 3837 CCTTACTCCAGCATTGCATAAACTACAGCTTGGTGAGGAGACTCAGTCAGATGAACCAAA 3896
Qy 592 ----------GlnSerLeuProLysLys------LysLysLysGlnLysLeuGluVa 605
Db 3897 CTTAGAGAGTCTGCAGTCTGAACCAAGAGAATTACCTCAAAGAAGTCAGGAGGCAAATAT 3956
Qy 605 lThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAl 625
Db 3957 GACAGAGAGCAGGAAG----------GC 3974
Qy 625 aLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIl 645
Db 3975 TGAAGATGAAATGCAGAAGTCAGCTTGGGATCAACCTTCACTTCCTGAAGGAAACAAA-- 4032
Qy 645 eGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAs 665
Db 4033 ----AATAAAACCAACTTGGATGAC---------CTAGTAAAGGGGGAAAATAGATCTTC 4079
Qy 665 pTyr----------AlaSerSerValPheAspThrAsn----- 674
Db 4080 AGTTAAACACAGATTGGCAGCCATGTCTAAAGCCAGCAGAAAATTCCCAGCTAAAGATGT 4139
Qy 675 -SerGlnGlnLysSerLeuAlaSer---------GlnSerThrGlnLysGluLeuGlnGl 691
Db 4140 AAGCCCCAGAAGACATGTAGCTACTATCTTCCCCCAAAGTGGAAGCAGATCTGGCTTTGA 4199
Qy 691 yHisLeuAlaLeuThrThrGlnGlu----------SerProHisProGlnAs 705
Db 4200 CCATTTATCTCTTGGCACAGTGGAGTGCAACCCACTGTTCCCTGAGCCTACTCCAAAATC 4259
Qy 705 nPheGlnSerThrGlnGluGlnGln----------ThrHisLeuArgMetGluGl 720
Db 4260 TGCAGAGTCCATTGGCGAAAGCAGGTTGAGTGAGAACGGAAAGCATGTGAAGAAATCCGA 4319
Qy 720 u----------MetValThrIleAl 725
Db 4320 GAACCTTCTCCCCATTACTGTACTACCCAACAGAGAACCTTCTACACACGTCAGCAACCA 4379
Qy 725 aAlaSerSerProLeuPheSerHisHisAspAspGlnTyr------IleAlaGluAlaPr 743
Db 4380 GAAGTCTAACAGCATTTCACAACGACATCAGAATGAGTTTAAAAACGTCTCAGAATCACC 4439
Qy 743 oThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaTh 763
Db 4440 ATCAAAGCATGAGAATTCTAAAGAC----------GTCAC 4469
Qy 763 rThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaVa 783
Db 4470 AGCAGCTCAGAATTTAGTAAGAGAATCAGGAGCCCCA----------TCACCCAT 4514
Qy 783 lAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProVa 803
Db 4515 CACATTCACCAGCCTCAGGGAAGCAGAATTCTCTGACAATCAGAGGAGGCTGAGCCCTCC 4574
Qy 803 lIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPh 823
Db 4575 TTTT---CCACTGGAGCCT---GCACAGAAA---------- 4599
Qy 823 eProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGl 843
Db 4600 ----TCTAGAGTAAGCAGTCCACTGGCCAGTTTTCTGCAGCAACAAAGGAGTGCTTCATC 4655
Qy 843 nValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspPr 863
Db 4656 TCTG----------GAGTGGGAACCTGAGCCACACCTCTATCGT---------TC 4691
Qy 863 oSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGl 883
Db 4692 AAAGAGTTTAAAAAGCATTAATGTTCATGGCGATCTACTACGAAAAAGC---------- 4740
Qy 883 nLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGl 903
Db 4741 ----CATCCTCCAAAAGTCAGGGAGCGCCATTTTTCTGAAAGCACTTCTATT-------- 4788
Qy 903 ySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHi 923
Db 4789 -GACAATGCCCTGAGTCGACTGACCCTTGGGAATGAATTCTCTGTCAACAATGGGTACAG 4847
Qy 923 sAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGl 943
Db 4848 TCGAAGATTCAGATCTTTTTCTGAACTCCCCTCCTGTGAT------GGAAAT-------- 4893
Qy 943 nLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyVa 963
Db 4894 ----------GAAAGTTGGGCTTATCGCAGCGGGACAAAAACAGGTCCCAGGTCTGCAAT 4943
Qy 963 lSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSe 983
Db 4944 ATCTATA----------TACAGACCTATCGACTA 4967
Qy 983 rGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAs 1003
Db 4968 TGGGATCTTTGGGAAAGAACAACAGTTAGCTTTCTTAGAAAATGTAAAGAGGTCA----- 5022
Qy 1003 pPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetVa 1023
Db 5023 ----CTTACACAAGGAAGATTATGGAAACCAAGTTTTCTTAAGAACCCTGGC-------- 5070
Qy 1023 lHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
Db 5071 ----------TTCCTAAAAGATGATTTG----------AGGAACCCT 5097

```
RESULT 14
ACH89376
ID   ACH89376 standard; DNA; 2612 BP.
XX
AC   ACH89376;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human genome derived single exon probe #22571.
XX
KW   Human; probe; ss; gene expression; single exon probe; microarray;
KW   alternative splicing event; genomic alteration.
XX
OS   Homo sapiens.
XX
PN   US2003194704-A1.
XX
PD   16-OCT-2003.
XX
PF   03-APR-2002; 2002US-00029386.
XX
PR   03-APR-2002; 2002US-00029386.
XX
PA   (PENN/) PENN S G.
PA   (RANK/) RANK D R.
PA   (HANZ/) HANZEL D K.
XX
PI   Penn SG,  Rank DR,  Hanzel DK;
XX
DR   WPI; 2004-119264/12.
XX
PT   New human genome-derived single exon nucleic acid probes useful for human
PT   gene expression analysis, for identifying or characterizing alternative
PT   splicing events, for assessing genomic alterations or as tools for
PT   surveying tissues.
XX
PS   Claim 1; SEQ ID NO 22571; 80pp; English.
XX
CC   The invention relates to a nucleic acid probe for measuring human gene
CC   expression, comprising any of the 27,400 fully defined nucleotide
CC   sequences in the specification, or their complements or fragments, and
CC   encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC   fully defined in the specification. The probe is a single exon probe that
CC   hybridises under high stringency conditions to a nucleic acid molecule
CC   expressed in human cells or tissues. Also included are a spatially-
CC   addressable set of single exon nucleic acid probes for measuring human
CC   gene expression (comprising a plurality of single exon nucleic acid
CC   probes cited above, where each of the plurality of probes is separately
CC   and addressably isolatable or amplifiable from the plurality), a single
CC   exon microarray for measuring human gene expression, a method of
CC   measuring human gene expression, a vector comprising the single exon
CC   probe cited above, an ORF-encoded peptide comprising at least 8
CC   contiguous amino acids of any of the above- mentioned amino acid
CC   sequences (optionally with conservative amino acid substitutions), an
CC   isolated antibody that binds specifically to a peptide cited above,
CC   methods of selling and/or licensing single exon probes or microarrays to
CC   a customer desiring to measure gene expression, a method of providing
CC   human gene expression data by subscription, and a computer-readable
CC   storage medium which contains a database having a plurality of records
CC   (each record including data on the expression of a single exon probe
CC   cited above. The probe, methods and apparatus are useful in gene
CC   expression analysis. The probes may be used as tools for surveying
CC   tissues to detect the presence of expressed messages that contain their
CC   specific exon, or in constructing genome-derived single exon microarrays.
CC   In addition, the probes are used in identifying and characterising
CC   alternative splicing events, in detecting and characterising gross
CC   alterations in the genomic locus that includes their exon, in assessing
CC   smaller genomic alterations, in priming the synthesis of nucleic acids,
CC   or in expressing the ORF-encoded peptide. The present sequence is a human
CC   single exon probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ   Sequence 2612 BP; 1060 A; 394 C; 550 G; 608 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.33e-06      Length:          2612
Score:                    223.00        Matches:         172
Percent Similarity:       34.3%         Conservative:    108
Best Local Similarity:    21.1%         Mismatches:      312
Query Match:              4.0%          Indels:          225
DB:                       12            Gaps:            33

US-09-828-068-2 (1-1057) x ACH89376 (1-2612)

Qy        62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
Db        86 AAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTA 145

Qy        82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
Db       146 ---AACAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTGAAACAAAA 202

Qy       101 ----------AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
Db       203 GCACGAAAAGGAGAAAAACCTTGTGCTTTGGAAAAGAAGGATATTTCAAAGTCAGAAGCT 262

Qy       114 LysLeu---------- 115
Db       263 AAACTTTCAAGAAAACAGGTAGATAGTGAGCACATGCATCAGAATGTTCCAACAGAGGAA 322

Qy       116 ----------LysThrSerAspAsnGlyThrAlaPro 124
Db       323 CAAAGAACAAATAAAAGTACCGGTGGTGAACATAAGAAATCTGATAGAAAAGAAGAACCT 382

Qy       125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
Db       383 CAATATGAACCTGCC---------AACACTTCTGAAGATTTAGACATGGATATTGTGTCT 433

Qy       145 SerThrPheValProAlaSerValGlySerGln----------LysValSerPro 159
Db       434 ---------GTTCCTTCCTCAGTTCCAGAAGACATTTTTGAGAATCTTGAGACTGCTATG 484

Qy       160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
Db       485 GAAGTTCAGAGTTCAGTTGATCATCAAGGGGATGGCAGCAGTGGAACTGAACAAGAAGTG 544

Qy       180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn------GlyAlaAlaGlu 197
Db       545 GAGAGTTCATCTGTAAAATTAAATATTTCTTCAAAAGACAACAGAGGAGGTATTAAATCA 604

Qy       198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db       605 AAAACTACAGCTAAAGTAACAAAAGAATTA----------TATGTTAAACTC 646

Qy       218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
Db       647 ACTCCTGTTTCCCTTTCTAATTCCCCAATTAAAGGTGCTGATTGTCAGGAAGTTCCACAA 706

Qy       238 IleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSer 257
Db       707 --------------GATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAG 751

Qy       258 GluValValLeuLysArgAsnGluAspGluAsnGly---------LysThrGluGluThr 274
Db       752 AAATGTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCA 811

Qy       275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
Db       812 TTACTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTTGAGG 871

Qy       295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314
Db       872 CGACCGACA----------GAAACTAACCCTGTAACATCTAATTCA 907
```

```
Qy      315 CysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSer 334
Db      908 GATGAAGAA---TGTAATGAAACAGTTAAGGAG------AAACAAAAACTATCAGTTCCA 958

Qy      335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354
Db      959 GTGAGAAAAAAGGATAAGCGT---------------------------------------- 979

Qy      355 GlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal------Arg 372
Db      980 ------AATTCTTCTGACAGTGCTATAGATAATCCTAAGCCTAATAAATTGCCAAAATCT 1033

Qy      373 LeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg---------SerAsp 389
Db     1034 AAGCAATCAGAGACTGTGGATCAAAATTCAGATTCTGATGAAATGCTAGCAATCCTCAAA 1093

Qy      390 GluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProVal 409
Db     1094 GAGGTGAGCAGGATGAGTCACAGTTCTTCTTCAGATACTGATATTAATGAAATTCATACA 1153

Qy      410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429
Db     1154 AACCATAAGACTTTGTATGATTTAAAGACTCAG---------GCGGGGAAAGAT------ 1198

Qy      430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp--------- 446
Db     1199 ------GATAAAGGAAAAAGGAAACGAAAAAGTTCTACATCTGGCTCAGATTTTGATACT 1252

Qy      447 ------GlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg------------ 460
Db     1253 AAAAAGGGCAAATCAGCTAAGAGCTCTATAATTTCTAAAAAGAAACGACAAACCCAGTCT 1312

Qy      461 --------------------------ThrGlySerValHisHisThrValAlaHisPro 471
Db     1313 GAGTCTTCTAATTATGACTCAGAATTAGAAAAAGAGATAAAGAGCATGAGTAAAATTGGT 1372

Qy      472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSer---ThrGlnHisAspAsp 490
Db     1373 GCTGCCAGAACCACCAAAAAAAGAATTCCAAATACAAAAGATTTTGACTCTTCTGAAGAT 1432

Qy      491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr---AspValCysGln 509
Db     1433 GAGAAACACAGCAAAAAAGGAATGGATAATCAAGGGCACAAAAATTTGAAGACCTCACAA 1492

Qy      510 HisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529
Db     1493 GAAGGATCATCTGATGATGCTGAAAGA---AAACAAGAGAGAGAGACTTTCTCTTCAGCA 1549

Qy      530 LysGly-----------------------------------------------------Lys 532
Db     1550 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTAAGAGATCGACTTCCTAAGAAG 1609

Qy      533 ThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIle 552
Db     1610 CAGCAAGCAAGTGCTTCCACT------GATGGTGTCGATAAGCTTTCTGGGAAAGAGCAG 1663

Qy      553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHis 572
Db     1664 AGTTTTACTTCTTTGGAAGTTAGAAAAGTTGCTGAAACTAAAGAAAAG------AGCAAG 1717

Qy      573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592
Db     1718 CATCTCAAAACCAAAACATGTAAAAAAGTACAGGATGGCTTATCTGATATTGCAGAGAAA 1777

Qy      593 SerLeuProLysLysLys----------------------LysLysGlnLysLeuGlu 604
Db     1778 TTCCTAAAGAAAGACCAGAGCGATGAAACTTCTGAAGATGATAAAAAGCAGAGCAAAAAG 1837

Qy      605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db     1838 GGAACTGAAGAAAAAAAGAAA--------------CCTTCAGACTTTAAGAAAAAAGTA 1882

Qy      625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp--------- 641
Db     1883 ATTAAA------ATGGAACAACAGTATGAATCTTCATCTGATGGCACTGAAAAGTTACCT 1936

Qy      642 --------------------------------IleAsnArgIleGlnSerLysThrThr 650
Db     1937 GAGCGAGAAGAAATTTGTCATTTTCCTAAGGGCATAAAACAAATTAAGAATGGAACAACT 1996

Qy      651 AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVal 670
Db     1997 --------------------------------GATGGAGAAAAGAAAAGTAAAAAAATA 2023

Qy      671 PheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690
Db     2024 AGAGATAAAACTTCTAAAAAGAAGGATGAATTATCTGATTATGCTGAGAAGTCAA--CAG 2081

Qy      691 GlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGln 710
Db     2082 GGA------------------------------------------------------- 2084

Qy      711 GluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeu 730
Db     2085 ----------------------AAGGAGATAGTTGTGACTCTT----------------- 2105

Qy      731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TrpGlyArgLy 750
Db     2106 -----------------CAGAGGATAAAAAGAGTAAGAATGGAGCATATGGTAGAGAGA 2147

Qy      750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db     2148 AGAAAAGGTGCAAGTTGCTTGGAAAGAGTTCAAGGAAGAGACAAGA 2193
```

RESULT 15
ADI21446
ID ADI21446 standard; cDNA; 4115 BP.
XX
AC ADI21446;
XX
DT 15-APR-2004 (first entry)
XX
DE Novel human expressed sequence tag, EST #145.
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;
KW ss; gene; EST; expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO2003025148-A2.
XX
PD 27-MAR-2003.
XX
PF 19-SEP-2002; 2002WO-US029964.
XX
PR 19-SEP-2001; 2001US-0323739P.
PR 13-SEP-2002; 2002US-00323739.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX
DR WPI; 2003-354603/33.
DR P-PSDB; ADI21666.
XX
PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.

XX
PS Example 2; SEQ ID NO 697; 156pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers,
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents a novel human expressed sequence tag, EST.
XX
SQ Sequence 4115 BP; 1590 A; 1191 C; 789 G; 545 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.42e-06	Length:	4115
Score:	223.00	Matches:	153
Percent Similarity:	35.5%	Conservative:	126
Best Local Similarity:	19.4%	Mismatches:	310
Query Match:	4.0%	Indels:	198
DB:	10	Gaps:	38

US-09-828-068-2 (1-1057) x ADI21446 (1-4115)

```
Qy        113 AspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsn 132
              |||:::         ::::::      |||   |||                 :::|||
Db       1623 GACAGGACTCCTTTGGCCAATGAGAAGACCACGCCATCTCTAGCAGAGCCTACAGAAAAT 1682

Qy        133 GlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerVal 152
              |||                              |||:::   |||
Db       1683 GGACAA--------------------------AGGACCCCATTTGCCAATGAGAAGACC 1715

Qy        153 GlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSer 172
               |||         |||:::          :::          |||:::   :::
Db       1716 ACATCATCCTCAGCAGAGCCTACAGAACACGAAGAAAGGACTCCACTGGCCAATGAGAAC 1775

Qy        173 ThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLys 192
              |||   |||   |||      |||      :::::::::                 :::
Db       1776 ACCACACCA---TCCCCGGCAGAGCCTACAGAAAATAGA-----------------GAA 1814

Qy        193 AsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGln 212
                  |||   |||       |||   ||||||                ||||||   :::
Db       1815 AGGACAGCCAATGAGAACACCACACCATCCCCA-----------GCAGGGCCTACAGAA 1862

Qy        213 AsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeu 232
              |||   ::::::   ||||||         :::   |||::::::
Db       1863 AATAGAGAAATGACAGCCAAC---------GAGAAGACCACACTATTCCCAGCAGAGCCT 1913

Qy        233 ProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSer 252
               |||      :::   |||  :::   :::          :::          |||:::
Db       1914 ACAGAAAATAGAGAAAGGACAGCCAATGAGAAGACCACATCATCCCCAGCAGAGCCTACA 1973

Qy        253 --------------ThrProLysLeuSerGluValValLeuLysArgAsnGluAspGlu 267
                           ||||||      :::|||                :::   :::
Db       1974 GAAAATGGACAAAGGACCCCATTTGCCAATGAGAAAACCACATCATCCTCAGCAGAGCCT 2033

Qy        268 AsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro--- 286
                    |||   |||   :::|||   :::   |||   |||      |||
Db       2034 ACAGAACACGGAGAAAGGACCCCACTGGCCAATGAG---AACACCACACTATCCCCAGCA 2090

Qy        287 AsnProMetSerGlyLysGlu------------------------------------- 293
               |||       :::|||
Db       2091 GGGCCTACAGAAAATAGAGAAAGGACAGCCAATGAGAAGACCACACCATTCCCAGCAGAG 2150

Qy        294 ------------ArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspPro------ 307
                         |||::::::   |||   :::   |||   |||      |||
```

```
Db       2151 CCTACAGAAAATAGAGAAAGGACAGCCAATGAG---AACACCACACCATCCCCAGCACAG 2207

Qy        308 --------------------LysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
                                  |||::::::   :::|||
Db       2208 CCTACAGAAAATGGAGACAGGACTCCATTGGCCAATGAGAAGACCACACCATCTCTAGCA 2267

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
              ||||||   |||            ||||||:::      :::      ||||||
Db       2268 GAGCCTACAGAAAATGGA------AAAAGGACCCCATTTGCCAATGAGAAGACCACATCA 2321

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
                    :::      :::|||:::::::::
Db       2322 TCCTCAGCAGAGCCTACAGAACACGCAGAAAGG-------------------------- 2354

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
                                                      :::   :::    |||
Db       2355 ----------------------------------------ACTCCACTGGCCAATGAG 2372

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAla---------- 396
              |||                              :::||||||
Db       2373 AACACCACATCATCCCCAGCAGAGCCTACAGAAAATAGAGAAAGGACAGCCAATGAGAAG 2432

Qy        397 --------------AlaAspProCysGluAspAspArgSer--------------Thr 406
                           |||:::|||   |||:::      |||                |||
Db       2433 ACCACACAATTCCCAGCAGAGCCTACAGAAAATAGAGAAAGCACAGCCAATGAGAAGACC 2492

Qy        407 IleProValProMetGluValSerMetAspIleProValSerAsnHis------ThrVal 424
               |||   |||   |||               |||      |||      |||
Db       2493 ACACCATTCCCAGCAGAG--------------CCTACAGAAAATAGAGAATGGACAGCC 2537

Qy        425 GlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValVal 444
              |||:::          |||            |||:::
Db       2538 AATGAGAACACCACACTATCCCCAGCAGAGCCTACAGAACAT----------------- 2579

Qy        445 AspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerVal 464
              ::::::   :::    |||   |||          :::|||   |||   |||
Db       2580 GAAGAAATGACCCCATTGGCCAAT--------------GAGAAGACCACACTATCC--- 2621

Qy        465 HisHisThrValAlaHisProAlaGlyAsn--------------LeuSerAsnLysLys 479
                        |||   |||      |||                :::|||:::|||
Db       2622 ---------CCAGCAGAGCCTACAGAAAATGGAGAAAGGACCCCATTTACCAATGAGAAG 2672

Qy        480 ValThrProThrAlaSer------ThrGlnHisAspAspGluAsnAspThrGluAsnGly 497
              ||||||:::::::::      |||:::|||   :::                |||
Db       2673 ACCACACCATCCTCAGCAGAGCCTACAGAACATGGAGAAAGGACCCCACTGGCCAATGAG 2732

Qy        498 LeuAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGln 517
              :::   |||         :::   :::      :::|||         |||:::   :::
Db       2733 ATC---ACCACACCATCCCGAGCAGAGCCTACAGAACATGGAGAAAGGATAGCCAATGAG 2789

Qy        518 ArgCys--------------------SerSerLysGlyLysThrAlaGlyLeuSerLys 530
              :::                     :::      |||:::|||          :::::::::
Db       2790 AAGGCCACACCATCCCCAGCAAAGCCTACAGAACATGGAGAAACGACA---GTCAATGAG 2846

Qy        531 GlyLysThrHisSerAlaAlaSerThrLysTyrGlyGly-------------------- 543
                   |||   |||:::|||                |||
Db       2847 GACACCACACCATCCTCAGCAGAGCCTACAGAAAATGGAGAAAGGACCCCACTGGCCAAT 2906

Qy        544 ----------------------GluSerThrArgAsnGlyGlnAsnIleHisValLeu 555
                                   |||||||||   :::|||:::
Db       2907 GAGAACACCACAACATCCCCAACAGAGTCTACAGAACATGGAGAAAGGACAGCCAATGAG 2966

Qy        556 SerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys 575
                                  ||||||:::         :::   ||||||
Db       2967 AAGACCACACCATCCCCAGCAGAGCCTACAGAACATGGAGAAAGGACACCATCAGCCAAT 3026

Qy        576 --------------------ValSerProAlaGluHisAspIleGlnIleMetSerAsp 588
                                  |||   ||||||:::   ::::::
Db       3027 GAGAAGACCATACCATCTCCAGCAAAGCCTACAGAACACGAA---GAAATGACCCCATCG 3083

Qy        589 LeuHisGluGlnSerLeuProLysLysLysLys------------GlnLysLeuGlu 604
                 :::|||   :::   |||      |||                :::|||
Db       3084 GCCAATGAGAACACCACACCATCCCCAGTAAAGCCTACAGAACATGGAGAAAAGACTACA 3143
```

```
Qy        605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
              :::      ||||||   |||:::      :::   |||            |||
Db       3144 TTGGCCAATGAGAAGATCACACTATCCCCAGAAGGGCCT---------ACAGAACATGGA 3194

Qy        625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArg 644
              ||||||                :::   :::|||   :::
Db       3195 GCAAAAACTACGTCGGCCAATGAGAAGATCACACCATCCCTAGCAAAGCCTACAGAACAT 3254

Qy        645 IleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySer 664
                  :::||||||:::   :::|||   :::      :::|||   :::   :::
Db       3255 ---GGAGAAAGGACCACATCACCCAATGACAAGATCACCTCATCTGCAGCAGAGTCTACA 3311

Qy        665 AspTyrAlaSerSerValPheAspThrAsn---------------------------- 674
              ::::::                  |||
Db       3312 GAACATAGAGATAGGGCTACATCAGCCAATGTGATCACACCAGCCCCAGCAGAGCCTATA 3371

Qy        675 ------SerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHis 692
                  :::::::::   :::||||||   :::         :::
Db       3372 AAACATGCAAAAAGGACCACATTGGCCCATGAGAAGATGACACAA---------------- 3416

Qy        693 LeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGln 712
                :::|||   :::   :::   ||||||:::         ||||||   |||:::
Db       3417 ------GTCACAGAAAAGTCCACAGAACACCCAGAAAAGACCACGTCAACCACAGAGAAA 3470

Qy        713 GlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSer 732
                |||      |||      |||      |||:::   :::         :::   :::
Db       3471 ACCACA------AGAACCCCAGAAAAGCCTACGCTATACTCAGAGAAGACCATATGCACC 3524

Qy        733 HisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAla 752
                    :::      :::   |||   |||||||||:::   |||
Db       3525 AAAGGGAAAAACACACCAGTCCCAGAAAAGCCTACAGAAAACCTGGGGAAC--------- 3575

Qy        753 LysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsn------SerProAlaAla 770
                ||||||   |||      ||||||      ::::::      :::|||
Db       3576 ACCACACTGACCACTGAGACCATAAAAGCCCCAGTAAAGTCCACAGAAAACCCAGAAAAA 3635

Qy        771 ThrCysGlyAla------GlnPheArgPro--------------GlyIleGlnAlaVal 783
              |||                     :::|||               |||   :::::::::
Db       3636 ACAGCAGCAGTCACAAAGACTATAAAACCTTCAGTCAAGGTCACAGGAGACAAATCTCTC 3695

Qy        784 AspLeuThrSerThrHisVal 790
                ||||||:::|||:::
Db       3696 ACTACTACCTCTTCTCATCTA 3716
```

Search completed: March 7, 2006, 14:35:36
Job time : 1628 secs

GenCore version 5.1.7

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2006, 13:34:46 ; Search time 9614 Seconds
(without alignments)
5143.951 Million cell updates/sec

Title: US-09-828-068-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTNCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09828068/runat_06032006_163155_25163/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05h
-USER=US09828068_@CGN_1_1_5315_@runat_06032006_163155_25163 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		5495	99.4	3174	10	CL958119	CL958119 OsIFCC000
c 2		1404	25.4	884	10	AG906393	AG906393 Oryza sat
c 3		1179.5	21.3	904	10	CZ880884	CZ880884 OC__Ba028
c 4		1028	18.6	723	10	CW759415	CW759415 OG_BBa006
5		973	17.6	838	10	CW621556	CW621556 OP__Ba001
6		854.5	15.5	857	10	CZ694942	CZ694942 OC__Ba000
c 7		820	14.8	1025	9	BZ795917	BZ795917 PUFHM45TD
8		814.5	14.7	888	9	CC355521	CC355521 PUHPL93TB
9		773.5	14.0	739	10	CZ785763	CZ785763 OC__Ba014
10		748.5	13.5	687	10	CL602016	CL602016 OB__Ba007
11		716.5	13.0	885	10	CG091207	CG091207 PUJFK24TD
12		691.5	12.5	460	10	CZ187525	CZ187525 OA_BBa011
13		646.5	11.7	647	10	CL597614	CL597614 OB__Ba005
c 14		645.5	11.7	731	10	CL704355	CL704355 SP__Bb001
15		606.5	11.0	787	10	CG345663	CG345663 OGWLW77TH
c 16		605	10.9	384	10	CW686090	CW686090 OG_BBa004
17		567.5	10.3	999	9	BZ795914	BZ795914 PUFHM45TB
c 18		518.5	9.4	766	9	CC603916	CC603916 OGWHM55TV
c 19		497.5	9.0	587	10	CL703548	CL703548 SP__Bb000
20		491.5	8.9	641	9	BZ411511	BZ411511 OGACC43TM
c 21		465.5	8.4	849	9	BZ411504	BZ411504 OGACC43TC
c 22		451.5	8.2	389	10	CZ821320	CZ821320 OC__Ba019
23		414	7.5	523	3	BJ246359	BJ246359 BJ246359
24		409	7.4	469	10	AG213697	AG213697 Oryza sat
25		401.5	7.3	556	5	BU974305	BU974305 HB27I10r
26		389	7.0	593	3	BQ244810	BQ244810 TaE15035B
c 27		387.5	7.0	674	10	CW161591	CW161591 104_568_1
c 28		371.5	6.7	653	9	BZ311518	BZ311518 ic68h08.b
29		349	6.3	822	10	CG365016	CG365016 OGHAD70TV
30		330.5	6.0	718	10	CW161592	CW161592 104_568_1
31		330	6.0	995	10	CG451574	CG451574 OG8AD34TV
c 32		326.5	5.9	736	10	CW365639	CW365639 fsbb001f0
33		320	5.8	469	1	AL818058	AL818058 AL818058
34		310	5.6	734	9	BZ326387	BZ326387 ic68h08.g
c 35		302.5	5.5	753	9	CC160109	CC160109 ig06g11.b
c 36		289	5.2	792	9	BZ422300	BZ422300 id52a11.b
37		287	5.2	491	10	AG209720	AG209720 Oryza sat
c 38		286	5.2	905	10	CG365008	CG365008 OGHAD70TH
c 39		275	5.0	970	10	CG032005	CG032005 PUFTA25TD
40		270.5	4.9	782	9	CC824592	CC824592 ZMMBBb016
c 41		266.5	4.8	903	9	BZ704364	BZ704364 PUBLF78TD
c 42		259.5	4.7	635	9	BH880296	BH880296 ht52d11.b
c 43		251.5	4.6	688	9	BZ776689	BZ776689 ii04h10.g
44		245.5	4.4	600	9	CC603904	CC603904 OGWHM55TH
45		241.5	4.4	703	9	BZ776688	BZ776688 ii04h10.b

ALIGNMENTS

RESULT 1
CL958119
LOCUS CL958119 3174 bp DNA linear GSS 21-SEP-2004
DEFINITION OsIFCC000847 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL958119
VERSION CL958119.1 GI:52370980
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 3174)
AUTHORS Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES Location/Qualifiers
source 1. .3174

/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3174
Score:	5495.00	Matches:	1052
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	5
Query Match:	99.4%	Indels:	0
DB:	10	Gaps:	0

US-09-828-068-2 (1-1057) x CL958119 (1-3174)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 60

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 120

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 180

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 240

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 300

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 360

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 420

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 480

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 540

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 600

Qy        201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 660

Qy        221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 720

Qy        241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 780

Qy        261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 840

Qy        281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 900

Qy        301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCAATTTGACCAAAGATCCAAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 960

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1020

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1080

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1140

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1200

Qy        401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGATGATATAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTATC 1260

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAAAACAAGACAAAACGCAAATAC 1320

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1380

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1440

Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1500

Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1560

Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1620

Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1680

Qy        561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db       1681 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCGCTCGGCAAAGGTTTCTCCAGCTGAG 1740

Qy        581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 1800

Qy        601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 1860

Qy        621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 1920

Qy        641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
```

```
Db       1921 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 1980

Qy        661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680

Db       1981 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2040

Qy        681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700

Db       2041 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2100

Qy        701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720

Db       2101 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2160

Qy        721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740

Db       2161 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2220

Qy        741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760

Db       2221 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2280

Qy        761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780

Db       2281 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2340

Qy        781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800

Db       2341 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2400

Qy        801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820

Db       2401 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCTGTTAACCAGGTCCATGCA 2460

Qy        821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840

Db       2461 AGAAATTTTCCAAGCACAATAGCAACCATGGAAACGAGTAAGTTATGTGATCGGAGAAAT 2520

Qy        841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860

Db       2521 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2580

Qy        861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880

Db       2581 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2640

Qy        881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900

Db       2641 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2700

Qy        901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920

Db       2701 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCGCGGCAT 2760

Qy        921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940

Db       2761 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGATTTGGTGTGCTTGGCTCC 2820

Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960

Db       2821 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 2880

Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980

Db       2881 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 2940

Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000

Db       2941 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3000

Qy       1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020

Db       3001 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3060

Qy       1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040

Db       3061 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3120

Qy       1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057

Db       3121 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3171
```

RESULT 2
AG906393/c

```
LOCUS       AG906393                884 bp    DNA     linear   GSS 03-NOV-2004
DEFINITION  Oryza sativa (indica cultivar-group) genomic DNA, BAC end sequence,
            BAC clone:K0471H01_R, genomic survey sequence.
ACCESSION   AG906393
VERSION     AG906393.1  GI:55372775
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Katagiri,S., Wu,J., Ito,Y., Karasawa,W., Shibata,M., Kanamori,H.,
            Katayose,Y., Namiki,N., Matsumoto,T. and Sasaki,T.
  TITLE     End Sequencing and Chromosomal in silico Mapping of BAC Clones
            Derived from an indica Rice Cultivar, Kasalath
  JOURNAL   Breeding Science 54, 273-279 (2004)
REFERENCE   2  (bases 1 to 884)
  AUTHORS   Sasaki,T., Matsumoto,T. and Wu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-2004) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki, 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The orientation of the sequence is from SP6 side of the BAC clone.
FEATURES             Location/Qualifiers
     source          1. .884
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Kasalath"
                     /db_xref="taxon:39946"
                     /clone="K0471H01_R"
ORIGIN
```

Alignment Scores:

Pred. No.:	1.3e-123	Length:	884
Score:	1404.00	Matches:	280
Percent Similarity:	96.3%	Conservative:	4
Best Local Similarity:	94.9%	Mismatches:	11
Query Match:	25.4%	Indels:	2
DB:	10	Gaps:	0

US-09-828-068-2 (1-1057) x AG906393 (1-884)

```
Qy        597 LysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIle 616

Db        884 AAGAGAAAAAGACGCAAAACTGAAGG-TGTCGGTNAAACCAGGCCATGATAGATGC-CTT 827

Qy        617 ProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGlu 636

Db        826 CCCATGGATATTGTTGAACTGCAAGCTAAAACCCAGCATGAGAGGCAGCTTATGACTGAG 767

Qy        637 ThrAspCysSerAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysVal 656

Db        766 ACTGATTGTTCTGACATCAACCGTATTCAATCCAAGCCAACTGCTGATGATGATTGTGTN 707

Qy        657 IleValAlaAlaLysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGln 676

Db        706 ATAGTAGCTGCCAAGGATGGTTCAGATAATGCATCAAGTGTGTTTGACACTAATTCCCAA 647
```

```
Qy        677 GlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThr 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACC 587

Qy        697 ThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeu 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 ACACAAGAGTCTCCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTG 527

Qy        717 ArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAsp 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CGGATGGAAGAAATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGAT 467

Qy        737 GlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThr 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 CAGTATATTGCTGAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACG 407

Qy        757 TrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPhe 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 TGGGAGCAATTTAAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTT 347

Qy        777 ArgProGlyIleGlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsn 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 AGACCTGGTATCCAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAAT 287

Qy        797 TyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsn 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TATGCATCTCGCCAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAAC 227

Qy        817 GlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCys 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CAGGTCCATGCAAGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGT 167

Qy        837 AspArgArgAsnAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHis 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GATCGGAGAAATGCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCAT 107

Qy        857 LeuLeuArgMetMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSer 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CTTCTGAGAATGATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGC 47

Qy        877 ArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrAlaHis 891
              |||||||||||||||||||||||||||||||||||||||||||||
Db         46 AGGAACCAGATGGAGTCTCAACTTCATAATTCTCAGTATGCACAT 2
```

RESULT 3
CZ880884/c

```
LOCUS       CZ880884                904 bp    DNA     linear   GSS 27-JUL-2005
DEFINITION  OC__Ba0283G18.f OC__Ba Oryza coarctata genomic clone OC__Ba0283G18
            5', genomic survey sequence.
ACCESSION   CZ880884
VERSION     CZ880884.1  GI:71356640
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0283  row: G  column: 18
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0283G18"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
```

Alignment Scores:

Pred. No.:	4.69e-102	Length:	904
Score:	1179.50	Matches:	235
Percent Similarity:	78.6%	Conservative:	18
Best Local Similarity:	73.0%	Mismatches:	16
Query Match:	21.3%	Indels:	53
DB:	10	Gaps:	5

US-09-828-068-2 (1-1057) x CZ880884 (1-904)

```
Qy         70 LysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLysCysAsp 89
              |||||||||||||||   ||||||||||||||||||||   ||||||||||||||||||
Db        903 AAGAAGGATCCAAAACTTTGCTCTCTATCTCGGATTTTCCGTGACCAGAAAAAATGTGAT 844

Qy         90 GluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCysSer 109
              :::|||:::   ||||||||||||||||   ||||||||||||||||||||||||||||||
Db        843 GATCACCAAGTTAGTTCAAGCCCATTTTTTGTAGCAAAGTTTCGACGGTGGGATTGCTCT 784

Qy        110 LysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAla 129
              |||||||||||||||||||||||:::||||||||||||||||||||   |||
Db        783 AAGTGCTTGGATAAGTTGAAAACTTCAGATCATGGAACAGCTCCAAGAACACTTTCCGCA 724

Qy        130 LysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValPro 149
              ||||||:::||||||||||||||||||||||||||||||||||||||||||:::|||
Db        723 AAGCAGGACGGCACAAGTGATGGTTGCTCCATCACATTTGTTCGGAGCACTTTTTTGCCG 664

Qy        150 AlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAla 169
              |||:::|||||||:::   ||||||||||||||||||||||||||||:::|||
Db        663 GCTACTGTTGGTTCCAAAAGTGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAGTGCT 604

Qy        170 AspArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaPro 189
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        603 GATAGATCAACTCTTCCGAAGAGTGCGCAAGAAGGCAATGACTCCAAATGCAATGCGCCT 544

Qy        190 SerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLys---------- 205
              |||||||||||||||||||||||||:::|||||||||||||||||||
Db        543 TCTGGCAAGAATGGAGCTGCTGAGGCCAGTACTGATTCACCCATGAAAGGTATGGTAGAT 484

Qy        205 ---------------------------------------------------------- 205

Db        483 GTAGAGCCTTTTAAATTCCTAGTAAGATTTTATTTAAGGATAGAATAAGATGATGTTTGT 424

Qy        206 ----------AspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSer 221
                       ||||||   ||||||||||||||||||:::|||||||||:::|||
Db        423 GTGATTTTTTCAGATTTGCATGGGCCAGCCCAAAATTATGATATGGCAGCAAATATCTCC 364

Qy        222 GluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrpHis 241
              |||:::|||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        363 GAGAACAACACTTCTGTAGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTATATGGCAC 304

Qy        242 IleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeu 261
                 :::   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 AAAAAAGGAAATGGTGCCGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCTC 244

Qy        262 LysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsn 281
              ||||||||||||||||||||||||||||||||||||   |||
Db        243 AAACGAAATGAAGATGAAAATGGAAAAACTGAAGAGATTCTT---------- 202
```

```
Qy        282 LeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCys 301
                                                      :::|||||||||||||
Db        201 ----------------------------------------ATTGCTGAGCAGTGC 187

Qy        302 AsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn--- 320
              |||   |||  ||||||||||||:::|||||||||||||||:::|||:::||||||
Db        186 AATTTTACCAATGATCCGAAACCAATGTCTGGACAGAAATGTGATCAGGTCTGCAACAGT 127

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSer------LysArgLysThr 338
              |||||||||   ||||||||||||||||||||||||||||      ||||||||||||
Db        126 GAGCCATGT---GAAGTTGTTCTGAAAAGAAGCTCCAAATCTAAATGTAAGAGGAAGACG 70

Qy        339 AspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspVal 358
              ||||||||||||||||||||:::|||   ||||||||||||||||||||||||||||
Db         69 GATAAGAAGTTGATGAAGAAGAAGCAGCAAAGCAAGAAACGCACTGCCCAGGCTGATGAT 10

Qy        359 SerAsp 360
              ||||||
Db          9 TCAGAT 4
```

RESULT 4
CW759415/c

LOCUS	CW759415 723 bp DNA linear GSS 09-NOV-2004
DEFINITION	OG_BBa0067J13.f OG_BBa Oryza glaberrima genomic clone OG_BBa0067J13 5', genomic survey sequence.
ACCESSION	CW759415
VERSION	CW759415.1 GI:55603764
KEYWORDS	GSS.
SOURCE	Oryza glaberrima (African rice)
ORGANISM	Oryza glaberrima Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1 (bases 1 to 723)
AUTHORS	Kim,H., Yu,Y., Wissotski,M., Byrne,M., Stum,D., Smart,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE	OMAP
JOURNAL	Unpublished (2004)
COMMENT	Contact: Rod A. Wing Arizona Genomics Institute University of Arizona Forbes Building Room 303, Tucson, AZ 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: rwing@genome.arizona.edu PCR PRimers FORWARD: TAA TAC GAC TCA CTA TAG GG BACKWARD: CAC TCA TTA GGC ACC CCA Plate: 0067 row: J column: 13 Seq primer: TAA TAC GAC TCA CTA TAG GG Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .723
	/organism="Oryza glaberrima"
	/mol_type="genomic DNA"
	/db_xref="taxon:4538"
	/clone="OG_BBa0067J13"
	/tissue_type="young leaves"
	/lab_host="DH10B T1 phage resistant"
	/clone_lib="OG_BBa"
	/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Alignment Scores:

Pred. No.:	1.12e-87	Length:	723
Score:	1028.00	Matches:	208
Percent Similarity:	86.3%	Conservative:	0
Best Local Similarity:	86.3%	Mismatches:	3
Query Match:	18.6%	Indels:	31
DB:	10	Gaps:	1

US-09-828-068-2 (1-1057) x CW759415 (1-723)

```
Qy        132 AsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSer 151
              |||||||||||||||||||||||||||||||||||||  |||||||||  |||||||||
Db        722 AATGGCACAAGTGATGGTTGCTCCATCACATTTGTTTGGAGCACTTTTGGCCCTGCTAGT 663

Qy        152 ValGlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLys-AsnAlaAspAr 171
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        662 GTTGGTTCCCAAAAAGTGTCTCCTAGCACACAATCATCTCAAGGGAAGAAATGCTGATAG 603

Qy        171 gSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGl 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATCAACTCTTCCAAAGAGTGTGCAAGAAGGCAATGACTCCAAATGCAATGCGCCTTCTGG 543

Qy        191 yLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLys----------------- 205
              |||||||||||||||||||||||||||||||||||||||||
Db        542 CAAGAATGGAGCTGCTGAGGCCAATACTGATTCACCAATGAA-AGGTATGGTAGATGTAG 484

Qy        205 ------------------------------------------------------------ 205

Db        483 AGCCTTTCAAATTCCTAAGTAGGATTTTATTTAAGGTATAGAATAAACTAATGTTTGTGT 424

Qy        206 ----------AspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGl 222
                         ||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GATTTTCTCAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTCTCTGA 364

Qy        222 uAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrpHisIl 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGGCACAT 304

Qy        242 eGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeuLy 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 AGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTCCTCAA 244

Qy        262 sArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLe 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGCAATTT 184

Qy        282 uThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAs 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAGTGCAA 124

Qy        302 nLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluPr 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAATGAGCC 64

Qy        322 oCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLe 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAGAAGTT 4

Qy        342 u 342
              |
Db          3 G 3
```

RESULT 5
CW621556

LOCUS	CW621556 838 bp DNA linear GSS 22-OCT-2004
DEFINITION	OP__Ba0019J18.f OP__Ba Oryza punctata genomic clone OP__Ba0019J18 5', genomic survey sequence.
ACCESSION	CW621556
VERSION	CW621556.1 GI:54547814
KEYWORDS	GSS.
SOURCE	Oryza punctata
ORGANISM	Oryza punctata Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1 (bases 1 to 838)
AUTHORS	SanMiguel,P., Westerman,R., Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C., Wing,R. and Jackson,S.A.
TITLE	OMAP Project - Purdue University
JOURNAL	Unpublished (2004)

```
COMMENT     Contact: Scott A. Jackson
            Jackson Laboratory
            Purdue University
            915 W. State St., West Lafayette, IN 47907, USA
            Tel: 7654963621
            Fax: 7654967255
            Email: sjackson@purdue.edu
            Basecalling by phred version 0.020425.c. This sequence was derived
            from the raw sequence read by clipping with lucy version 1.19s.
            Bases 46-883 of the raw sequence (length 1374) were retained after
            clipping.
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 161   Std Error: 0.00
            Plate: 0019  row: J  column: 18
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Oryza punctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4537"
                     /clone="OP__Ba0019J18"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OP__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:              2.71e-82        Length:         838
Score:                  973.00          Matches:        182
Percent Similarity:     98.4%           Conservative:   3
Best Local Similarity:  96.8%           Mismatches:     3
Query Match:            17.6%           Indels:         0
DB:                     10              Gaps:           0

US-09-828-068-2 (1-1057) x CW621556 (1-838)

Qy        870 ProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyr 889
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||  ||||||
Db          2 CCCAACTATGTAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTTTCAGTAT 61

Qy        890 AlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLys 909
              |||   ||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db         62 GCACCTAATCAGTACAAAGGATCAACTAGCACATCATATGGCAGTAACATAAATGGAAAG 121

Qy        910 IleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeu 929
              |||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        122 ATTCCATTGACATTCGAAGACTTATCACGGCATCAGATGCATGATTTGCACAGACCTTTA 181

Qy        930 ArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrp 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAACTGG 241

Qy        950 SerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleThrSer 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCGGAGAACTGTGGCACACAATCTGGTTATAAGTTGGGAGTGTCGACAGGGATAACATCG 301

Qy        970 HisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaLys 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CATCAGATGAACAGAAAGGAACATTTTGAAGCCTTGAATTCTGGAATGTTTTCAGCAAAA 361

Qy        990 TrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaArgAsn 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAGGAAC 421

Qy       1010 SerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPhe 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATAGGTTT 481

Qy       1030 ValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsn 1049
              ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db        482 GTGAGACAGGATATCTGTGTAACTAACAAGAATCCAGCTGATTTTACTACAATCAGTAAC 541

Qy       1050 AspAsnGluTyrMetAspTyrArg 1057
              |||||||||||||||||||||||||
Db        542 GATAACGAGTATATGGATTACCGG 565

RESULT 6
CZ694942
LOCUS       CZ694942                857 bp    DNA     linear   GSS 25-JUL-2005
DEFINITION  OC__Ba0007O13.f OC__Ba Oryza coarctata genomic clone OC__Ba0007O13
            5', genomic survey sequence.
ACCESSION   CZ694942
VERSION     CZ694942.1  GI:71094090
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 857)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0007  row: O  column: 13
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .857
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0007O13"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:              6.59e-71        Length:         857
Score:                  854.50          Matches:        177
Percent Similarity:     70.3%           Conservative:   24
Best Local Similarity:  61.9%           Mismatches:     42
Query Match:            15.5%           Indels:         44
DB:                     10              Gaps:           3

US-09-828-068-2 (1-1057) x CZ694942 (1-857)

Qy        363 LeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGln 382
              ||||||  ||||||||||||||||||||||||||||||||||||||||  |||:::|||
Db          5 CTTTGTTGGAGAAAGCCGAAAAAGGTACGGCTTTTATCGGAAATTATAAAAGCTGACCAA 64

Qy        383 ValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAsp 402
              ||||||  |||||||||:::||||||  ||||||  ||||||||||||||||||||||||
Db         65 GGTGAGGATCCTAGAAGTGATAAAGTTCATCATGAAAATGTTGCTGATCCCTGTGAGGAT 124

Qy        403 AspArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHis 422
              :::||||||  |||  |||:::|||||||||||||||||||  |||:::
Db        125 GAGAGAAGCTTAATC------CCATTGGAAGTAAGCATGGATATTCCTGGTAGCCACCAG 178
```

```
Qy        423 ThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAsp 442
Db        179 AAAGTGGGAGATGATGGGTCAAAATCAACTAAGAACAAGACAGAACGCAAATGCACTGAT 238

Qy        443 ValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg---Thr 461
Db        239 GCTGTAGATGATGGATCATCACTCATGAACTGGCTGAATGGAAAGAAGAAAAAAGTAACT 298

Qy        462 GlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThr 481
Db        299 GAAAATGTGCATCACACTGCTGTACATCCAGCTGGGAATTTGAGCAATAGAAAAGTGACC 358

Qy        482 ProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsn 501
Db        359 CCCAGTGTCAGTGCTCAGCATGATGATGAGGATAATATTCAAAATGGTCTTGACATAAAT 418

Qy        502 MetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSer 521
Db        419 ATGCATAAGACTGATATCTGCCAACATGAATCAGAAAACTCTACACAGAGGTGTTCGTCA 478

Qy        522 LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyr 541
Db        479 AAGGGGAAAACAGCGGGTTTGAGTAAGTGGAAAACACATTCAACTGCCAGTGCCAAAAAT 538

Qy        542 GlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCys 561
Db        539 GGTGATGAAAACACCAGAAACAGTCAGAATATACCTATACCCAGGACAGAAGATCAATGC 598

Qy        562 GlnMetGluThrGluAsnSerVal---------------------------------- 569
Db        599 CAGGTGGAAACCGAAAACTCTGTTCTGAGGTGCTTGACAAAGGTAAGAATTTTGTGTTTT 658

Qy        569 ------------------------------------------------------------ 569
Db        659 TCTTTTATACTATGGTGCTATATTTTTGTTCCGACTGAGATTATAGTGGTTATTTTCCTA 718

Qy        570 ------------------------LeuSerHisSerAlaLysValSerProAlaGluHis 581
Db        719 TATTGTAATGGACTGACATTTTTCCTCAT-TTTTCATGTCAGGTTTCTCCAGCTAAGCGT 777

Qy        582 AspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGln 601
Db        778 GATATTCAAAATGTGGTTAGCCTTCATGAGCAGAGTCTTCCCAAGAAGAAAAAGAAGCCA 837

Qy        602 LysLeuGluValThrArg 607
Db        838 AATCTTGAACTGGTGCGT 855
```

RESULT 7
BZ795917/c

```
LOCUS       BZ795917               1025 bp    DNA     linear   GSS 17-MAR-2003
DEFINITION  PUFHM45TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa320H17,
            genomic survey sequence.
ACCESSION   BZ795917
VERSION     BZ795917.1  GI:28996773
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUFHM45TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .1025
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa320H17"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN
```

Alignment Scores:

Pred. No.:	1.8e-67	Length:	1025
Score:	820.00	Matches:	179
Percent Similarity:	64.6%	Conservative:	40
Best Local Similarity:	52.8%	Mismatches:	94
Query Match:	14.8%	Indels:	26
DB:	9	Gaps:	9

US-09-828-068-2 (1-1057) x BZ795917 (1-1025)

```
Qy        678 LysSerLeuAlaSerGlnSerThrGln---LysGluLeuGlnGlyHis---LeuAlaLeu 695
Db        982 AAATGTTCAGCCTCCAGAATCACAAATGCAAAATCGCTTCAGGTTCATGTGTTACCCATC 923

Qy        696 ThrThrGlnGluSerProHisProGlnAsn--------------------PheGlnSer 708
Db        922 AAACGGTCTTTCATTGTGTACCCTCAAAAATTACCTGTCCTTGATGACATTTTGGAGTGT 863

Qy        709 ThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSer 728
Db        862 ACCCAAGAACACCAGACACAATTCCACAGGGATGGAGGGGTCACCATTGCATGTACCTCA 803

Qy        729 ProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGly 748
Db        802 CCTATGTTTTCACATCAT------CAACATATTGCTGAAGTGCCTGCTCGGAGCTGGAGG 749

Qy        749 ArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPro 768
Db        748 AATAACGGGGAAAAGAAGTTAATGTGGGATTCTTTCAAGACAGCTTCAAGAAATTCACCA 689

Qy        769 AlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThr 788
Db        688 ACATCGTCATATGGTTTTCAATTCGGAAACAGGCTTCAAGAAGTTCATTCATCTCCCAGT 629

Qy        789 HisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAsp 808
Db        628 CATGCTTATGGAGCTTCTAGTAACTATGCAGCTCACCAGCCAGTAATTGTGGCTGTAGAT 569

Qy        809 ArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAla 828
Db        568 CAG----------------------------CCGAGAAGTGTTCCAAGCACAACATCG 539

Qy        829 ThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPro 848
Db        538 ACTATGGAGGCTGGTAGGTTGTATGATCGAAGGATTCCTGGACAGTCAGGCCTCTACCCA 479

Qy        849 LysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSer 868
Db        478 AAAGAGCCTATGCCTGCAACACATCTTCTGAGATTGATGGACTCATCAACAGCTCCAGAC 419

Qy        869 PheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHis------Asn 886
Db        418 TTCACAAACTATCAAAGAGCTAACAGGCGCCAGATGGAACATCAAACACAAACTCTGGAT 359

Qy        887 SerGlnTyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsn 905
Db        358 TCACAATATACACAGCATGATCAGTATAATGGGTCACTGAGCACATCATATGGAAGACAC 299
```

```
Qy        906 LeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeu---HisAsp 924
              :::   ||||||:::|||   |||   :::||||||||||||   |||:::      :::
Db        298 ATAACTGGAAAGGTTCCACCGACACTGCAAGACTTATCTCGGCGTCAGGTCCAGCAAAAC 239

Qy        925 LeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLys 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        238 CTGCACAGGCCTTTACGCCCTCATCCTCGGGTGGGCGTGCTTGGTTCGTTGCTGCAGCAG 179

Qy        945 GluIleAlaAsnTrpSerGluAsnCysGly---ThrGlnSerGlyTyrLysLeuGlyVal 963
              :::||||||||||||   ||||||||||||      ||||||||||||:::||||||||
Db        178 GATATCGCAAACTGGTATGAAAACTGCGGGCCGCCACAGTCTGGATACAGGCTGGGCGTT 119

Qy        964 SerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSer 983
              |||   |||      |||      |||||||||      ::::::|||   |||||||||
Db        118 TCTAAAGGGACGGCATCGCTTGATATGAACAGACATGGAAACTACGAGACCTTGAACTCA 59

Qy        984 GlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla 1002
              |||:::|||:::|||   ||||||||||||||||||||||||||||||||   |||
Db         58 GGAGTGTTCACAGCAGGATGGAATGCCCTTCAATTGGGTTCTGTTAGCTCTCTTGCC 2

RESULT 8
CC355521
LOCUS       CC355521                888 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  PUHPL93TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa504O18,
            genomic survey sequence.
ACCESSION   CC355521
VERSION     CC355521.1  GI:30824921
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHPL93TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa504O18"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Alignment Scores:
Pred. No.:              4.81e-67        Length:         888
Score:                  814.50          Matches:        169
Percent Similarity:     69.8%           Conservative:   37
Best Local Similarity:  57.3%           Mismatches:     72
Query Match:            14.7%           Indels:         17
DB:                     9               Gaps:           6

US-09-828-068-2 (1-1057) x CC355521 (1-888)

Qy        767 SerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThr 786
              ||||||   ::::::   |||   ||||||         :::|||   |||      :::
```

```
Db          2 TCACCAACATCGTCATATGGTTTTCAATTCGGAAACAGGCTTCAAGAAGTTCATTCATCT 61

Qy        787 SerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaPro 806
              :::|||      |||:::|||||||||||||||:::   ||||||||||||
Db         62 CCCAGTCATGCTTATGGAGCTTCTAGTAACTATGCAGCTCACCAGCCAGTAATTGTGGCT 121

Qy        807 LeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnProSerThr 826
              :::|||:::                                  |||:::   |||||||||
Db        122 ATAGATCAG----------------------------CCGAGAAGTGTTCCAAGCACA 151

Qy        827 IleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeu 846
              :::|||||||||||   :::|||   ||||||||      ||||||      |||
Db        152 ACATCGACTATGGAGGCTGGTAGGTTGTATGATCGAAGGATTCCTGGACAGTCAGGCCTC 211

Qy        847 TyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeu 866
              ||||||||||||   ||||||||||||||||||||||||:::||||||   ||||||
Db        212 TACCCAAAAGAGCCTATGCCTGCAACACATCTTCTGAGATTGATGGACTCATCAACAGCT 271

Qy        867 AlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHis--- 885
                    |||   ||||||      ::::::|||   |||||||||   |||
Db        272 CCAGACTTCACAAACTATCAAAGAGCTAACAGGCGCCAGATGGAACATCAAACACAAACT 331

Qy        886 ---AsnSerGlnTyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGly 903
                 :::|||||||||      |||:::||||||   ||||||   ||||||||||||||||
Db        332 CTGGATTCACAATATACACAGCATGATCAGTATAATGGGTCACTGAGCACATCATATGGA 391

Qy        904 SerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeu--- 922
              ::::::   ||||||:::|||   |||   :::||||||||||||   |||:::
Db        392 AGACACATAACTGGAAAGGTTCCACCGACACTGCAAGACTTATCTCGGCGTCAGGTCCAG 451

Qy        923 HisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeu 942
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 CAAAACCTGCACAGGCCTTTACGCCCTCATCCTCGGGTGGGCGTGCTTGGTTCGTTGCTG 511

Qy        943 GlnLysGluIleAlaAsnTrpSerGluAsnCysGly---ThrGlnSerGlyTyrLysLeu 961
              |||::::::||||||||||||   ||||||||||||      ||||||||||||:::|||
Db        512 CAGCAGGATATCGCAAACTGGTATGAAAACTGCGGGCCGCCACAGTCTGGATACAGGCTG 571

Qy        962 GlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeu 981
              |||||||||   |||      |||      |||||||||      ::::::|||   |||
Db        572 GGCGTTTCTAAAGGGACGGCATCGCTTGATATGAACAGACATGGAAACTACGAGACCTTG 631

Qy        982 AsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSer 1001
              |||||||||:::|||:::|||   ||||||||||||||||||||||||||||||||||||
Db        632 AACTCAGGAGTGTTCACAGCAGGATGGAATGCCCTTCAATTGGGTTCTGTTAGCTCTCTT 691

Qy       1002 AlaAsp------PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLys 1019
              ||||||      :::         |||             |||      |||   |||
Db        692 GCCGATCCAGAGTACCGGTTGCCAAGGTATGGTACAGGTCAGCCTTGTACAGGTGGCAAT 751

Qy       1020 GlyLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLys 1039
              |||:::   ||||||||||||||:::   |||||:::|||||||||:::||||||:::
Db        752 GGGAGAACGGTTCACCCGTTGGATAAGCTTGTGCGAAAGGATATCTGTGTGACTAACAGA 811

Qy       1040 AsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
              ||||||||||||||||||   ||||||:::   ||||||||||||
Db        812 AACCCAGCTGATTTTACTGTAATTAGTGACATGAATGAGTATATG 856

RESULT 9
CZ785763
LOCUS       CZ785763                739 bp    DNA     linear   GSS 26-JUL-2005
DEFINITION  OC__Ba0148N23.f OC__Ba Oryza coarctata genomic clone OC__Ba0148N23
            5', genomic survey sequence.
ACCESSION   CZ785763
VERSION     CZ785763.1  GI:71225616
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
```

```
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0148  row: N  column: 23
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0148N23"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:                3.07e-63        Length:         739
Score:                    773.50          Matches:        154
Percent Similarity:       81.7%           Conservative:   20
Best Local Similarity:    72.3%           Mismatches:     35
Query Match:              14.0%           Indels:         4
DB:                       10              Gaps:           2

US-09-828-068-2 (1-1057) x CZ785763 (1-739)

Qy        366 ArgLysProLysLysVal-ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAs 385
              |||||||||||||||||| |||||||||||||||||||   |||:::|||   |||||
Db          2 AGAAAGCCGAAAAAGGTACAGGCTTTTATCGGAAATTATAAAAGCTGACCAAGGTGAGGA 61

Qy        385 pSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSe 405
              |    ||||||||:::||||||   ||||||   ||||||||||||||||||:::||||||
Db         62 TCCTAGAAGTGATAAAGTTCATCATGAAAATGTTGCTGATCCCTGTGAGGATGAGAGAAG 121

Qy        405 rThrIleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGl 425
              |   |||      |||:::|||||||||||||||||||   |||:::      |||||
Db        122 CTTAATC------CCATTGGAAGTAAGCATGGATATTCCTGGTAGCCACCAGAAAGTGGG 175

Qy        425 yGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAs 445
              |:::||||||   ||||||:::|||||||||||:::||||||   :::|||   |||||
Db        176 AGATGATGGGTCAAAATCAACTAAGAACAAGACAGAACGCAAATGCACTGATGCTGTAGA 235

Qy        445 pAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg---ThrGlySerVa 464
              |||||||||||||||||||||||||||||||||||||||||:::   |||   :::||
Db        236 TGATGGATCATCACTCATGAACTGGCTGAATGGAAAGAAGAAAAAAGTAACTGAAAATGT 295

Qy        464 lHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAl 484
              ||||||||||      |||||||||||||||||||||:::|||||||||||:::
Db        296 GCATCACACTGCTGTACATCCAGCTGGGAATTTGAGCAATAGAAAAGTGACCCCCAGTGT 355

Qy        484 aSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLy 504
                |||   ||||||||||||||:::::   :::||||||||||||   ||||||||||
Db        356 CAGTGCTCAGCATGATGATGAGGATAATATTCAAAATGGTCTTGACATAAATATGCATAA 415

Qy        504 sThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLy 524
              |||||||:::|||||||||   ||||||   ||||||||||||||||||||||||||||
Db        416 GACTGATATCTGCCAACATGAATCAGAAAACTCTACACAGAGGTGTTCGTCAAAGGGGAA 475

Qy        524 sThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGl 544
              |||||||||||||||||||   ||||||||||||   ||||||   |||   |||   ||
Db        476 AACAGCGGGTTTGAGTAAGTGGAAAACACATTCAACTGCCAGTGCCAAAAATGGTGATGA 535

Qy        544 uSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGl 564
              |:::|||||||||   |||||||||   :::          |||||||||||||||:::||
Db        536 AAACACCAGAAACAGTCAGAATATACCTATACCCAGGACAGAAGATCAATGCCAGGTGGA 595

Qy        564 uThrGluAsnSerValLeuSerHisSerAlaLysVal 576
              ||||||||||||||||||||              ||||||
Db        596 AACCGAAAACTCTGTTCTGAGGTGCTTGACAAAGGTA 632

RESULT 10
CL602016
LOCUS       CL602016                687 bp    DNA     linear   GSS 10-JUN-2004
DEFINITION  OB__Ba0072E10.r OB__Ba Oryza brachyantha genomic clone
            OB__Ba0072E10 3', genomic survey sequence.
ACCESSION   CL602016
VERSION     CL602016.1  GI:48579779
KEYWORDS    GSS.
SOURCE      Oryza brachyantha
  ORGANISM  Oryza brachyantha
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE     OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 145000   Std Error: 0.00
            Plate: 0072  row: E  column: 10
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Oryza brachyantha"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4533"
                     /clone="OB__Ba0072E10"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OB__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:                6.82e-61        Length:         687
Score:                    748.50          Matches:        152
Percent Similarity:       79.5%           Conservative:   19
Best Local Similarity:    70.7%           Mismatches:     41
Query Match:              13.5%           Indels:         3
DB:                       10              Gaps:           2

US-09-828-068-2 (1-1057) x CL602016 (1-687)

Qy        364 CysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnVal 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTCGGAGAAAGCCGAAAAAGGTACGGCTTCTATCAGAAATTATAAATGCTAACCAAGGT 60

Qy        384 GluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAsp 403
                  :::   ||||||||||||   |||   ||||||   :::|||       ||||||:::
Db         61 GGGGAACCTAGAAGCGATGAAGATCATCATGAAAATGTTTCTGATCACCGTGAGGATGAG 120
```

```
Qy        404 ArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHisThr 423
              |||     |||      |||:::|||||||||||||||    |||   |||:::
Db        121 AGAAGGTTAATC------CCATTGGAAGTAAGCATGGATTTTCCTGGTAGCCACCAGAAA 174

Qy        424 ValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspVal 443
              |||   ||||||  |||||||||:::|||:::|||||||||||||||:::::|||
Db        175 GTAAGAGAAGATGAATTAAAGTCAACTAAGAGCAAGACAAAACGCAAATTCGCTGATGCT 234

Qy        444 ValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLys---LysArgThrGly 462
              |||||||||||||||||||||||||||||||||||||||||||   |||   |||
Db        235 GTAGATGATGGATCATCACTCATGAACTGGCTGAATGGAAAAAAGAGAAAAACAACTGAA 294

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
              :::|||||||||:::|||      |||||||||||||||||||:::||||||||||||
Db        295 AATGTGCATCACAGTGTTGTACGTCCAGCTGGGAATTTGAGCAACAGAAAAGTGACCCCC 354

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMet 502
              |||   |||   |||||||||||:::::::::   ||||||||||||||   ||||||
Db        355 ACTGTCAGTGCTCAGCATGATGATGACGATAATATTGAAAATGGTCTTGACGTAAACATG 414

Qy        503 HisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLys 522
              |||   ||||||:::  ||||||   ||||||   ||||||||||||||||||||||
Db        415 CATATGACAGATATGCGCCAACATGAATCAGAGAACTCTACACAGAGGTGCTCTTCAAAG 474

Qy        523 GlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGly 542
              |||   ||||||||||||||||||   ||||||||||||   ||||||   |||   |||
Db        475 GGAACAACGGCGGGTTTGAGTAAGTGGAAAACACATTCACCTGCCAGTGCCAAAAATGGT 534

Qy        543 GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGln 562
              |||:::||||||:::   |||||||||   :::|||      ||||||||||||||||
Db        535 GATGAAAACACTAGAGACAGTCAGAACATACCTATACTCAGGACAGAAGATCAATGCCAG 594

Qy        563 MetGluThrGluAsnSerValLeuSerHisSerAlaLysValSer 577
              ||||||||||||||||||||||      ||||||||||||||||
Db        595 ATGGAAACCGAAAACTCTGTTCTGAGGTGCTCGGCAAAGGTAAGC 639

RESULT 11
CG091207
LOCUS       CG091207                 885 bp    DNA     linear   GSS 20-AUG-2003
DEFINITION  PUJFK24TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0656C23,
            genomic survey sequence.
ACCESSION   CG091207
VERSION     CG091207.1  GI:33973501
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUJFK24TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa0656C23"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

Alignment Scores:
Pred. No.:              1.21e-57        Length:         885
Score:                  716.50          Matches:        160
Percent Similarity:     61.0%           Conservative:   29
Best Local Similarity:  51.6%           Mismatches:     88
Query Match:            13.0%           Indels:         33
DB:                     10              Gaps:           9

US-09-828-068-2 (1-1057) x CG091207 (1-885)

Qy        753 LysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsn------SerProAlaAla 770
              |||||||||   |||:::   ||||||      :::||||||      ||||||
Db         27 AAGAAGCTAATGTGGGATTCTTTCAAGACAGCTTCAAGGAATCCACCATCACCACCACCG 86

Qy        771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
              :::   |||   ||||||          :::|||   ||||||              |||
Db         87 TCATATGGTTTTCAATTCGGAAACAGGCTTCAAGAAGTTGATTCAGCTCCCATTCATGCA 146

Qy        791 MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyr 810
                                  |||   :::   ||||||||||||      :::|||:::
Db        147 -----------------TATGGAGCTCACCAGCCAGTAATTGTGGCCGTAGATCAG--- 185

Qy        811 AlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMet 830
                                                |||:::   |||||||||      |||
Db        186 --------------------------CCGAGAAGTGTTCCAAGCACAGCGTTGGCTATG 218

Qy        831 GluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProLysGlu 850
              ||||||   :::|||   |||:::|||      ||||||      ||||||||||||||
Db        219 GAGGCTGGTAGGCTGTACGATCAAAGGATCCCTGGACAGCCAGGCCTTTACCCAAAAGAG 278

Qy        851 SerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro 870
              |||||||||||||||||||||||:::||||||   ||||||         |||
Db        279 CCCATGCCTGCAACACATCTCCTGAGGCTGATGGACTCATCAACAGCTCCAGGCTTCACA 338

Qy        871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu------SerGlnLeuHisAsnSerGln 888
              ||||||      :::::::|||   |||||||||      :::|||         ||||||
Db        339 AACTACCAAAGAGCTAATAGGCGGCAGATGGAGCTTCAAACGCAAGCTCTCGGCTCGCAG 398

Qy        889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907
              |||   |||   |||   |||   ||||||||||||||||||
Db        399 TACACACAGCATGAGCATTACAATGCGTCACCGAGCACGTCATATGGAAGCCAA------ 452

Qy        908 GlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926
              |||||||||   ||||||||||||   ||||||         ||||||
Db        453 ---------CCGCTGACACTGGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGCGC 503

Qy        927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        504 AGGCCTTTACGCCCTCACCCTCGTGTCGGTGTCTCGGTTCATTGCTGCAGCAGGAGATC 563

Qy        947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGly 966
              ||||||||||||   |||||||||   ||||||||||:::||||||          |||
Db        564 GCAAACTGGTCTGGGAACTGCGGGCCGCAGTCTGGGTACAGACTAGGTGATTGTAAAGGG 623

Qy        967 IleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPhe 986
                                  ||||||      :::::|||   |||:::
Db        624 ACAAGGCCGCCTGATGCCAACAGAGCAGGAAACTACGAGACCTTGAGC------------ 671

Qy        987 SerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla------AspPhe 1004
              ||||||  ||||||  |||||||||||||||||||||:::|||      ::::::
Db        672 TCGGCAGGATGGAACACCCTGCAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGTAC 731

Qy       1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024
                                  |||   |||   |||   |||:::   :::|||
Db        732 CGGTTTCCATGGTATGGTACGGGTCAGCCTTCGACCAGTGGCAATGGGAGGACCATTCAT 791

Qy       1025 ProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPhe 1044
              |||||||||:::   ||||||:::|||||||||:::||||||:::|||||||||||||||
Db        792 CCGTTGGATAAGCTCGTGAGGAAGGATATCTGTGTGACTAACAGAAACCCAGCCGATTTC 851
```

```
Qy       1045 ThrThrIleSerAsnAspAsnGluTyrMet 1054
              |||   ||||||:::   ||||||||||||
Db        852 ACTGTAATTAGTGACAAGAACGAGTACATG 881

RESULT 12
CZ187525
LOCUS       CZ187525                460 bp    DNA     linear   GSS 03-FEB-2005
DEFINITION  OA_BBa0113K02.r OA_BBa Oryza alta genomic clone OA_BBa0113K02 3',
            genomic survey sequence.
ACCESSION   CZ187525
VERSION     CZ187525.1  GI:58544739
KEYWORDS    GSS.
SOURCE      Oryza alta
  ORGANISM  Oryza alta
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   SanMiguel,P., Westerman,R., Kim,H., Yu,Y., Wissotski,M., Yost,D.,
            Stum,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C.,
            Hatfield,J., Soderlund,C., Wing,R. and Jackson,S.A.
  TITLE     OMAP (Oryza Map Alignment Project)- Purdue University
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Scott A. Jackson
            Jackson Laboratory
            Purdue University
            915 W. State St., West Lafayette, IN 47907, USA
            Tel: 7654963621
            Fax: 7654967255
            Email: sjackson@purdue.edu
            Basecalling by phred version 0.020425.c. This sequence was derived
            from the raw sequence read by clipping with lucy version 1.19s.
            Bases 96-555 of the raw sequence (length 1027) were retained after
            clipping.
            Plate: 0113  row: K  column: 02
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Oryza alta"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:52545"
                     /clone="OA_BBa0113K02"
                     /tissue_type="young leaves"
                     /lab_host="DH10B-T1 phage resistant"
                     /clone_lib="OA_BBa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:              1.05e-55        Length:         460
Score:                  691.50          Matches:        140
Percent Similarity:     91.0%           Conservative:   2
Best Local Similarity:  89.7%           Mismatches:     10
Query Match:            12.5%           Indels:         4
DB:                     10              Gaps:           2

US-09-828-068-2 (1-1057) x CZ187525 (1-460)

Qy        364 CysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnVal 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db          1 TGTCGGAGAAAGCCAAAAAAGGTCCGGCTTCTATCAGAAATTATAAACACTAACCAGGTT 60

Qy        384 GluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAsp 403
              |||||||||||||||||:::||||||   ||||||   |||||||||   |||||||||
Db         61 GAGGATTCTAGAAGCGACAAAGTTCATCATGAAAATGTCGCTGATCCCTATGAGGATGAT 120

Qy        404 ArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHisThr 423
              ||||||||||||||||||      ||||||||||||||||||||||||   |||   |||
Db        121 AGAAGTACCATCCCGGTG------GAAGTAAGCATGGATATTCCTGTTAGAAACCAGACA 174

Qy        424 ValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspVal 443
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        175 GTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATGCTCTGATGTT 234

Qy        444 ValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg---ThrGly 462
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        235 GTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGAACAACTGGA 294

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
              :::||||||||||||||||   ||||||||||||||||||||||||   ||||||||||||
Db        295 AATGTGCATCACACAGTTGTGCATCCAGCTGGGAATTTGAGCAATATAAAAGTGACACCC 354

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMet 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACAAACATG 414

Qy        503 HisLysThrAspValCys-GlnHisValSerGluIleSerThrGln 517
              ||||||||||||||||   ||||||||||||||||||||||||||||
Db        415 CATAAGACAGATGTCTGGTCAGCATGTATCAGAAATCTCAACACAG 460

RESULT 13
CL597614
LOCUS       CL597614                647 bp    DNA     linear   GSS 10-JUN-2004
DEFINITION  OB__Ba0059J21.r OB__Ba Oryza brachyantha genomic clone
            OB__Ba0059J21 3', genomic survey sequence.
ACCESSION   CL597614
VERSION     CL597614.1  GI:48570923
KEYWORDS    GSS.
SOURCE      Oryza brachyantha
  ORGANISM  Oryza brachyantha
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
            Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
  TITLE     OMAP Project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Insert Length: 145000   Std Error: 0.00
            Plate: 0059  row: J  column: 21
            Seq primer: CAC TCA TTA GGC ACC CCA
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Oryza brachyantha"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4533"
                     /clone="OB__Ba0059J21"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OB__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:              3.79e-51        Length:         647
Score:                  646.50          Matches:        137
Percent Similarity:     66.7%           Conservative:   21
Best Local Similarity:  57.8%           Mismatches:     27
Query Match:            11.7%           Indels:         53
DB:                     10              Gaps:           4

US-09-828-068-2 (1-1057) x CL597614 (1-647)
```

```
Qy        119 AspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCys 138
Db          3 GATTATGGAACATCACCAAGAACTCTTCCTGCAAAGCAAGATGGCACAAGCGATGGTTGC 62

Qy        139 SerIleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer 158
Db         63 TCCATCACATTTGTTCGGAGCACTTTTATGCCTGCTACTGTTGGTTCGAAAATAGTGTCT 122

Qy        159 ProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerVal 178
Db        123 CCTAGTGCACAATCATCTCAAGGGAAGAAGTCTGATAGATCAACTCTTCCGAAGAGTGCG 182

Qy        179 GlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAla 198
Db        183 CAAGAAGGCAACAACTCCAAATGCAATGCACCTTCTGGCCAGAAAGGAGCTGCTGAGGGC 242

Qy        199 AsnThrAspSerProMetLys---------------------------------------- 205
Db        243 AATACTGG-TCCACACATGAAAGGTATGGTCGAGGTAGAGCTTTTTAAATTCCTAGATAG 301

Qy        206 ----------------------------------------------AspLeuGlnGly 209
Db        302 GTTTTTATTTAAGGTATAGAATAAACTGATGTTTGTGTGATTCCCACAGATTTGCAAGGG 361

Qy        210 ProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspVal 229
Db        362 GCAGCCCAAAATTATGATATGGCAGTAAATATC---CCTGACAACACTTCTATAGATGTT 418

Qy        230 GlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGln 249
Db        419 GGAGTGTTGCCTGAAGTTCCCCAGATTGCATTGCACAAAGAAGGGAACGATGAAGATCAA 478

Qy        250 ProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGly 269
Db        479 TCCCCATCCACTCCAAAACTTTCTGAAGTGATCCTCAAACCAAATGATGATAGAGATGGA 538

Qy        270 LysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMet 289
Db        539 AAAACTAAAGGACTTCTT---------------------------------------- 556

Qy        290 SerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysPro 309
Db        557 --------------------GTTGCTGAGAAGTGCAATTTGACCAAACATCTGAAACCA 595

Qy        310 ValSerGlyGlnLysCysGluGlnIleCysAsn---GluProCysGluGlu 325
Db        596 ATGTCTGGACAGAAATGCGATCAGGTCTGCAACAGTGAGCCATGTGAAGAC 646
```

RESULT 14
CL704355/c

```
LOCUS       CL704355                731 bp    DNA     linear   GSS 22-NOV-2004
DEFINITION  SP__Bb0014N03.r SP__Bb Sorghum propinquum genomic clone
            SP__Bb0014N03 3', genomic survey sequence.
ACCESSION   CL704355
VERSION     CL704355.2  GI:55940476
KEYWORDS    GSS.
SOURCE      Sorghum propinquum
  ORGANISM  Sorghum propinquum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Kudrna,D.,
            Muller,C., Soderlund,C., Bowers,J.E., Paterson,A.H. and Wing,R.
  TITLE     Physical mapping of the sorghum genome
  JOURNAL   Unpublished (2004)
COMMENT     On Nov 22, 2004 this sequence version replaced gi:50270630.
            Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            Plate: 0014  row: N  column: 03
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Sorghum propinquum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:132711"
                     /clone="SP__Bb0014N03"
                     /lab_host="DH10B"
                     /clone_lib="SP__Bb"
                     /note="Vector: pCUGI1; Site_1: HindIII; Site_2: HindIII;
                     NSF Sorghum propinquum library"
ORIGIN
```

Alignment Scores:

Pred. No.:	5.77e-51	Length:	731
Score:	645.50	Matches:	136
Percent Similarity:	70.2%	Conservative:	34
Best Local Similarity:	56.2%	Mismatches:	65
Query Match:	11.7%	Indels:	7
DB:	10	Gaps:	4

US-09-828-068-2 (1-1057) x CL704355 (1-731)

```
Qy        792 GlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAla 811
Db        728 GGAGCCTCTAGTAACTATGCAGCTCACCAGCCAGTAATTGTGGCTGTAGATCCCTACACA 669

Qy        812 GluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGlu 831
Db        668 AAGGAAGCAGTTAGCCTGGTTCATNCAAGAAATGTCCCAAGCACAGCATTGACTATGGGG 609

Qy        832 AlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProLysGluSer 851
Db        608 TCTGGTAGGCTGTACGATCAAAGGATTCCTGGACAGTCAGGCCTTTATCCAAAAGAGCCT 549

Qy        852 MetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheProAsn 871
Db        548 GTGCCCGCTACACATCTTCTGAGACTGATGGACTCATCAACACCTTCAGGCTTCACAAGC 489

Qy        872 TyrGlyThrSerSerArgAsnGlnMetGlu---------SerGlnLeuHisAsnSerGln 888
Db        488 TATCAAAGACCCGACAGGCGCCAGATGGAACTTCATCAAACACAAACTCTCGGTTCACAA 429

Qy        889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907
Db        428 TTTACGCAGCATGATCAGTATAATGCATCACCGAGCACATCATATGGAAGCCACATAATT 369

Qy        908 GlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHis---AspLeuHis 926
Db        368 GAAGAGGTTCCACTGACACTGCAGGACTTATCTCGGCGTCAGCTCCAGCCAAACTTGCTC 309

Qy        927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946
Db        308 AGGCCTGTACGCCCTCATCCTCGAGTTGGCGTGCTTGGTTCATTGCTGCAGCAGGATATT 249

Qy        947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGly 966
Db        248 GCAAACTGGTCTGAAAACTATGGGCCACAGTCTGGGTACAGATTAGGTGATTCTAAAGGG 189

Qy        967 IleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPhe 986
Db        188 ACAACAACGCTTGATATGAACAGAAACGGAAACTACAAGGCCTTGAACTCAGGAATGCTC 129

Qy        987 SerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla------AspPhe 1004
Db        128 TCAGCAGGATGGAGTGCTCTGCAGTTGGGTTCTGTTACCTCTGTTGCCAATCCAGAGTAC 69

Qy       1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024
Db         68 CAATTGCCAAGGTATGGTACAGGTCAACCTTCTACCAGTGGCAATGGGAAAACAATTCAT 9
```

```
Qy       1025 ProLeu 1026
              ||||||
Db          8 CCGTTG 3

RESULT 15
CG345663
LOCUS       CG345663                787 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OGWLW77TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0614M09,
            genomic survey sequence.
ACCESSION   CG345663
VERSION     CG345663.1  GI:34262929
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .787
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0614M09"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

Alignment Scores:
Pred. No.:              3.59e-47        Length:         787
Score:                  606.50          Matches:        144
Percent Similarity:     58.9%           Conservative:   25
Best Local Similarity:  50.2%           Mismatches:     84
Query Match:            11.0%           Indels:         34
DB:                     10              Gaps:           9

US-09-828-068-2 (1-1057) x CG345663 (1-787)

Qy        760 PheLysAlaThrThrArgAsn------SerProAlaAlaThrCysGlyAlaGlnPheArg 777
              ||||||      :::|||||||      ||||||      :::   |||   ||||||
Db          3 TTCAAGACAGCTTCAAGGAATCCACCATCACCACCACCGTCATATGGTTTTCAATTCGGA 62

Qy        778 ProGlyIleGlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyr 797
                  :::|||   ||||||            |||                      |||
Db         63 AACAGGCTTCAAGAAGTTGATTCAGCTCCCATTCATGCA-----------------TAT 104

Qy        798 AlaSer-ArgGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGl 817
                 |||    ||||||||||||      :::|||:::
Db        105 AGGAGCTCACCAGCCAGTAATTGTGGCCGTAGATCAG---------------------- 141

Qy        817 nValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAs 837
                  |||:::   |||||||||          |||||||||   :::|||   ||
Db        142 -------CCGAGAAGTGTTCCAAGCACAGCGTTGGCTATGGAGGCTGGTAGGCTGTACGA 194

Qy        837 pArgArgAsnAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLe 857
              |:::|||      ||||||      ||||||||||||||   ||||||||||||||||||
Db        195 TCAAAGGATCCCTGGACAGCCAGGCCTTTACCCAAAAGAGCCCATGCCTGCAACACATCT 254

Qy        857 uLeuArgMetMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerAr 877
              |||||||:::||||||   ||||||          |||   ||||||      ::::::||
Db        255 CCTGAGGCTGATGGACTCATCAACAGCTCCAGGCTTCACAAACTACCAAAGAGCTAATAG 314

Qy        877 gAsnGlnMetGlu------SerGlnLeuHisAsnSerGlnTyrAla---HisAsnGlnTy 894
              |    |||||||||      :::|||          |||||||||      |||      ||
Db        315 GCGGCAGATGGAGCTTCAAACGCAAGCTCTCGGCTCGCAGTACACACAGCATGAGCATTA 374

Qy        894 rLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPh 914
              |      |||   ||||||||||||||||||                  |||||||||
Db        375 CAATGCGTCACCGAGCACGTCATATGGAAGCCAA--------------CCGCTGACACT 419

Qy        914 eGluAspLeuSer---ArgHisGlnLeuHisAspLeuHisArgProLeuArgProHisPr 933
              ||||||||||||   ||||||          ||||||   ||||||||||||||||||||
Db        420 GGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGCGCAGGCCTTTACGCCCTCACCC 479

Qy        933 oArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCy 953
              |||||||||||||||||||||||||||||:::|||||||||||||||||   |||||
Db        480 TCGTGTCGGTGTGCTCGGTTCATTGCTGCAGCAGGAGATCGCAAACTGGTCTGGGAACTG 539

Qy        953 sGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleThrSerHisGlnMetAs 973
              ||||   ||||||||||||:::||||||          |||                ||
Db        540 CGGGCCGCAGTCTGGGTACAGACTAGGTGATTGTAAAGGGACAAGGCCGCCTGATGCCAA 599

Qy        973 nArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLe 993
              ||||       ::::::|||   |||:::            ||||||  ||||||   ||
Db        600 CAGAGCAGGAAACTACGAGACCTTGAGC-----------TCGGCAGGATGGAACACCCT 647

Qy        993 uGlnLeuGlySerValSerSerSerAla------AspPheLeuSerAlaArgAsnSerIl 1011
              |||||||||||||||||||:::|||      ::::::
Db        648 GCAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGTACCGGTTTCCATGGTATGGTAC 707

Qy       1011 eAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPheValAr 1031
              |||      |||   |||   |||:::   :::||||||||||:::   |||||
Db        708 GGGTCAGCCTTCGACCAGTGGCAATGGGAGGACCATTCATCCGTTGGATAAGCTCGTGAG 767

Qy       1031 gGlnAspIleCysIleThr 1037
              |:::||||||||:::|||
Db        768 GAAGGATATCTGTGTGACT 786


Search completed: March  7, 2006, 16:32:05
Job time : 9638 secs
```

GenCore version 5.1.7

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2006, 13:51:31 ; Search time 438 Seconds
(without alignments)
4289.686 Million cell updates/sec

Title: US-09-828-068-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTNCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09828068/runat_06032006_163158_25224/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss05h -USER=US09828068_@CGN_1_1_193_@runat_06032006_163158_25224
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	298	5.4	8648	3	US-09-415-946-2	Sequence 2, Appli
	2	298	5.4	17341	3	US-09-415-946-1	Sequence 1, Appli
	3	219	4.0	6745	3	US-09-774-528-286	Sequence 286, App
	4	219	4.0	6745	3	US-10-120-988-286	Sequence 286, App
	5	213	3.9	8733	3	US-10-144-198-11	Sequence 11, Appl
	6	209	3.8	6852	3	US-10-172-502-3	Sequence 3, Appli
	7	199.5	3.6	6414	3	US-09-134-001C-1626	Sequence 1626, Ap
c	8	197.5	3.6	3561	3	US-09-134-001C-1685	Sequence 1685, Ap
	9	196.5	3.6	7571	3	US-09-949-016-4366	Sequence 4366, Ap
	10	196.5	3.6	8146	3	US-09-976-594-725	Sequence 725, App
	11	196.5	3.6	11917	3	US-09-566-921-32	Sequence 32, Appl
	12	196.5	3.6	101300	3	US-09-949-016-16108	Sequence 16108, A
	13	196.5	3.6	145287	3	US-09-949-016-13530	Sequence 13530, A
	14	196.5	3.6	145287	3	US-09-949-016-13531	Sequence 13531, A
	15	196	3.5	6755	3	US-08-931-999-4	Sequence 4, Appli
	16	194.5	3.5	6775	3	US-09-620-312D-289	Sequence 289, App
	17	192	3.5	9700	3	US-09-698-295-9	Sequence 9, Appli
	18	192	3.5	9865	3	US-09-698-295-2	Sequence 2, Appli
	19	190.5	3.4	4832	3	US-09-949-016-1478	Sequence 1478, Ap
	20	190.5	3.4	4832	3	US-09-949-016-1479	Sequence 1479, Ap
	21	189	3.4	5992	3	US-09-949-016-546	Sequence 546, App
	22	188	3.4	5992	3	US-09-949-016-2263	Sequence 2263, Ap
	23	188	3.4	7518	3	US-09-620-312D-1051	Sequence 1051, Ap
c	24	186.5	3.4	2835	3	US-09-134-001C-1515	Sequence 1515, Ap
	25	185	3.3	8538	3	US-09-774-528-98	Sequence 98, Appl
	26	185	3.3	8538	3	US-10-120-988-98	Sequence 98, Appl
	27	184.5	3.3	5709	2	US-08-425-061-8	Sequence 8, Appli
	28	184.5	3.3	5709	2	US-08-825-886-8	Sequence 8, Appli
	29	184.5	3.3	5709	3	US-08-989-890-8	Sequence 8, Appli
	30	184.5	3.3	8503	3	US-09-620-312D-130	Sequence 130, App
	31	183	3.3	2628	3	US-09-614-221A-444	Sequence 444, App
	32	183	3.3	5707	2	US-08-425-061-11	Sequence 11, Appl
	33	183	3.3	5707	2	US-08-825-886-11	Sequence 11, Appl
	34	183	3.3	5707	3	US-08-989-890-11	Sequence 11, Appl
	35	183	3.3	5711	2	US-08-425-061-10	Sequence 10, Appl
	36	183	3.3	5711	2	US-08-825-886-10	Sequence 10, Appl
	37	183	3.3	5711	2	US-08-658-322-1	Sequence 1, Appli
	38	183	3.3	5711	3	US-08-989-890-10	Sequence 10, Appl
	39	183	3.3	5711	3	US-09-462-401A-1	Sequence 1, Appli
	40	183	3.3	5712	2	US-08-425-061-12	Sequence 12, Appl
	41	183	3.3	5712	2	US-08-825-886-12	Sequence 12, Appl
	42	183	3.3	5712	2	US-08-603-753D-1	Sequence 1, Appli
	43	183	3.3	5712	3	US-09-099-753-1	Sequence 1, Appli
	44	183	3.3	5712	3	US-08-986-106-1	Sequence 1, Appli
	45	183	3.3	5712	3	US-09-007-678B-47	Sequence 47, Appl

ALIGNMENTS

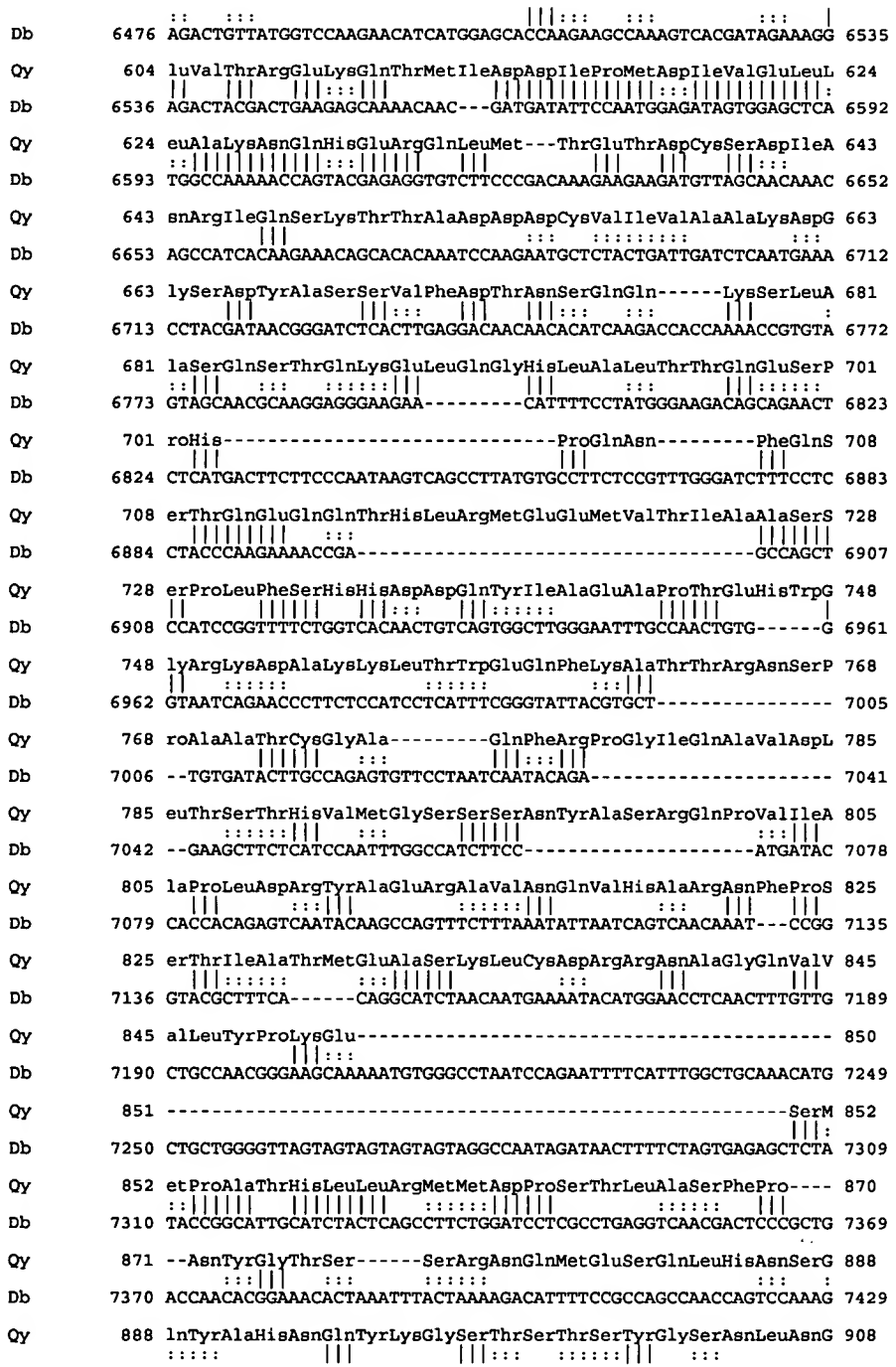
RESULT 1
US-09-415-946-2
; Sequence 2, Application US/09415946
; Patent No. 6376751
; GENERAL INFORMATION:
; APPLICANT: Sung, Z. Renee
; APPLICANT: Aubert, Dominique
; APPLICANT: Chen, Lingjing
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Reproductive Development in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 018941-000110US
; CURRENT APPLICATION NUMBER: US/09/415,946
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,696
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 8648
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: EMF1 gene and its promoter region
; NAME/KEY: promoter
; LOCATION: (1)..(3201)
; NAME/KEY: CDS
; LOCATION: join(4241..4335, 4448..4623, 4704..4823, 4903..4956,
; OTHER INFORMATION: EMBRYONIC FLOWER 1 (EMF1)
; NAME/KEY: exon
; LOCATION: (3202)..(3265)
; OTHER INFORMATION: exon 1
; NAME/KEY: intron

```
;   LOCATION: (3266)..(4159)
;   NAME/KEY: exon
;   LOCATION: (4160)..(4335)
;   OTHER INFORMATION: exon 2
;   NAME/KEY: intron
;   LOCATION: (4336)..(4447)
;   NAME/KEY: exon
;   LOCATION: (4448)..(4623)
;   OTHER INFORMATION: exon 3
;   NAME/KEY: intron
;   LOCATION: (4624)..(4703)
;   NAME/KEY: exon
;   LOCATION: (4704)..(4823)
;   OTHER INFORMATION: exon 4
;   NAME/KEY: intron
;   LOCATION: (4824)..(4902)
;   NAME/KEY: exon
;   LOCATION: (4903)..(4956)
;   OTHER INFORMATION: exon 5
;   NAME/KEY: intron
;   LOCATION: (4957)..(5045)
;   NAME/KEY: exon
;   LOCATION: (5046)..(6307)
;   OTHER INFORMATION: exon 6
;   NAME/KEY: intron
;   LOCATION: (6308)..(6447)
;   NAME/KEY: exon
;   LOCATION: (6448)..(8065)
;   OTHER INFORMATION: exon 7
;   NAME/KEY: intron
;   LOCATION: (8066)..(8300)
;   NAME/KEY: exon
;   LOCATION: (8301)..(8648)
;   OTHER INFORMATION: exon 8
US-09-415-946-2

Alignment Scores:
Pred. No.:                1.09e-17      Length:          8648
Score:                    298.00        Matches:         225
Percent Similarity:       35.5%         Conservative:    152
Best Local Similarity:    21.2%         Mismatches:      360
Query Match:              5.4%          Indels:          324
DB:                       3             Gaps:            50

US-09-828-068-2 (1-1057) x US-09-415-946-2 (1-8648)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
              :::|||      |||      :::   |||      |||   |||         |||
Db       5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
                |||   |||::::::                |||:::|||:::   :::      |||
Db       5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
                   :::         :::         |||:::
Db       5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
                |||   ||||||          |||      :::   :::          |||
Db       5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy        267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
              :::||||||   :::|||        :::|||
Db       5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT--------------------------- 5353

Qy        287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
                  |||||||||  :::|||             :::|||                 |||
Db       5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy        304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
                  :::   |||        |||          :::
Db       5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
              ||||||      ||||||:::   |||         ||||||     :::|||||:::
Db       5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy        344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
                 :::      :::   |||:::      :::                |||:::
Db       5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy        361 AlaLysLeuCys---------------------------------------------- 364
              :::|||   |||
Db       5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy        365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
                           :::::::::   ::::::   :::::::::    |||   :::
Db       5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy        380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
              :::         |||   |||:::      |||      :::|||:::   |||||||||
Db       5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy        400 CysGluAspAspArgSerThr--------------------------------IlePro 408
              :::         |||||||||                                :::|||
Db       5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy        409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
                 |||                  |||   :::      |||      |||
Db       5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy        429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
              |||         |||   |||||||||:::           |||:::|||:::|||
Db       5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy        449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
              |||   :::            |||   ::::::|||   :::            ||||||
Db       5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
                  :::||||||:::                |||:::   |||:::|||
Db       5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
                       :::         :::              |||
Db       5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy        500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
                          |||::::::|||      |||                |||    :::
Db       6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy        516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
                   |||                    |||||||||     :::   |||   |||
Db       6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy        529 SerLysGly---------------------------------------------LysThr 533
                    |||                                              :::
Db       6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy        534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
                 |||              :::|||      ||||||
Db       6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy        554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
              ||||||              |||         :::      ::::::||||||||||||
Db       6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy        573 r------------------------------------------------------------ 573
              |
Db       6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy        574 --------------------------Ala-LysValSerProAlaGluHisAspIleG 584
                                         |||         |||   |||:::      :
Db       6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475
```

```
Qy       584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuG 604
Db      6476 AGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGATAGAAAGG 6535

Qy       604 luValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuL 624
Db      6536 AGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTGGAGCTCA 6592

Qy       624 euAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSerAspIleA 643
Db      6593 TGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGCAACAAAC 6652

Qy       643 snArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspG 663
Db      6653 AGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTCAATGAAA 6712

Qy       663 lySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------LysSerLeuA 681
Db      6713 CCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAACCGTGTA 6772

Qy       681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerP 701
Db      6773 GTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAGCAGAACT 6823

Qy       701 roHis----------------------------ProGlnAsn---------PheGlnS 708
Db      6824 CTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATCTTTCCTC 6883

Qy       708 erThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerS 728
Db      6884 CTACCCAAGAAAACCGA---------------------------------GCCAGCT 6907

Qy       728 erProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpG 748
Db      6908 CCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG------G 6961

Qy       748 lyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerP 768
Db      6962 GTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT---------------- 7005

Qy       768 roAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAlaValAspL 785
Db      7006 --TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA---------------------- 7041

Qy       785 euThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleA 805
Db      7042 --GAAGCTTCTCATCCAATTTGGCCATCTTCC--------------------ATGATAC 7078

Qy       805 laProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProS 825
Db      7079 CACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT---CCGG 7135

Qy       825 erThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValV 845
Db      7136 GTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAACTTTGTTG 7189

Qy       845 alLeuTyrProLysGlu---------------------------------------- 850
Db      7190 CTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGCAAACATG 7249

Qy       851 ----------------------------------------------------SerM 852
Db      7250 CTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAGAGCTCTA 7309

Qy       852 etProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhePro---- 870
Db      7310 TACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACTCCCGCTG 7369

Qy       871 --AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHisAsnSerG 888
Db      7370 ACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAGTCCAAAG 7429

Qy       888 lnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnG 908
Db      7430 AGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG------- 7482

Qy       908 lyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgP 928
Db      7483 --CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCTTCCCGGA 7534

Qy       928 roLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleA 947
Db      7535 AGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTTCAAAATG 7588

Qy       947 laAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyI 967
Db      7589 CTTCATGGAGTCCTCAT---------------------------------------- 7605

Qy       967 leThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheS 987
Db      7606 --------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC--------------TTTG 7642

Qy       987 erAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp---------- 1003
Db      7643 CTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAAGCGAAGT 7702

Qy      1004 -----PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysM 1022
Db      7703 TCCAGCTGCTGGGAGCATCGAATTCC--------------------------------A 7729

Qy      1022 etValHisProLeu---------------AspArgPheValArgGlnAsp---------- 1033
Db      7730 TGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGAAAAGCAG 7789

Qy      1034 ----------------------------------------------------------IleC 1035
Db      7790 AGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCCATTGTGT 7849

Qy      1035 ysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMet 1054
Db      7850 GCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTTTACATG 7908
```

RESULT 2
US-09-415-946-1
; Sequence 1, Application US/09415946
; Patent No. 6376751
; GENERAL INFORMATION:
; APPLICANT: Sung, Z. Renee
; APPLICANT: Aubert, Dominique
; APPLICANT: Chen, Lingjing
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Reproductive Development in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 018941-000110US
; CURRENT APPLICATION NUMBER: US/09/415,946
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,696
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 17341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: genomic DNA (Ecotype Columbia) from CD82 clone
; NAME/KEY: promoter
; LOCATION: (1)..(3201)
; NAME/KEY: CDS
; LOCATION: join(4241..4335, 4448..4623, 4704..4823, 4903..4956,
; OTHER INFORMATION: EMBRYONIC FLOWER 1 (EMF1)
; NAME/KEY: exon
; LOCATION: (3202)..(3265)
; OTHER INFORMATION: exon 1
; NAME/KEY: intron
; LOCATION: (3266)..(4159)

```
;   NAME/KEY: exon
;   LOCATION: (4160)..(4335)
;   OTHER INFORMATION: exon 2
;   NAME/KEY: intron
;   LOCATION: (4336)..(4447)
;   NAME/KEY: exon
;   LOCATION: (4448)..(4623)
;   OTHER INFORMATION: exon 3
;   NAME/KEY: intron
;   LOCATION: (4624)..(4703)
;   NAME/KEY: exon
;   LOCATION: (4704)..(4823)
;   OTHER INFORMATION: exon 4
;   NAME/KEY: intron
;   LOCATION: (4824)..(4902)
;   NAME/KEY: exon
;   LOCATION: (4903)..(4956)
;   OTHER INFORMATION: exon 5
;   NAME/KEY: intron
;   LOCATION: (4957)..(5045)
;   NAME/KEY: exon
;   LOCATION: (5046)..(6307)
;   OTHER INFORMATION: exon 6
;   NAME/KEY: intron
;   LOCATION: (6308)..(6447)
;   NAME/KEY: exon
;   LOCATION: (6448)..(8065)
;   OTHER INFORMATION: exon 7
;   NAME/KEY: intron
;   LOCATION: (8066)..(8300)
;   NAME/KEY: exon
;   LOCATION: (8301)..(8648)
;   OTHER INFORMATION: exon 8
US-09-415-946-1

Alignment Scores:
Pred. No.:              3.69e-17        Length:         17341
Score:                  298.00          Matches:        225
Percent Similarity:     35.5%           Conservative:   152
Best Local Similarity:  21.2%           Mismatches:     360
Query Match:            5.4%            Indels:         324
DB:                     3               Gaps:           50

US-09-828-068-2 (1-1057) x US-09-415-946-1 (1-17341)

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db       5090 TCTAATGTTAGGAGCAAGAGCAGAAAGCTTGTGAGTCCGGAGCAG---------GTAGGA 5140

Qy        218 AlaAsnValSerGluAsp---------AsnThrSerValAspValGlyAlaLeuProGlu 234
Db       5141 AACAACAGATCTAAAGAAAAACTAAACAAACCATCGATGGATATTAGCAGCTGGAAAGAG 5200

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAla---------------------- 247
Db       5201 AAACAAAATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT 5260

Qy        248 ---AspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAsp 266
Db       5261 GAGGATACACCACCTAAGGCAACCAAGAATCATAAAGGCATTCGCGGTCTGATGGAATGC 5320

Qy        267 GluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspPro 286
Db       5321 GATAACGGGTCATCAGAAAGTATAAATCTTGCT---------------------------- 5353

Qy        287 AsnProMetSerGlyLysGluArg---------AspGlnValAlaGluGlnCysAsnLeu 303
Db       5354 ------ATGAGTGGGTTGCAGCGTAGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTG 5407

Qy        304 ThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCys 323
Db       5408 CTTGGTAATACAAAAACCAGTGGTGGTAGTAACATCAGAAAA----------------- 5449

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db       5450 GAAGAGTCTGCTTTGAAGAAGGAATCAGTTAGAGGTCGAAAA------AGAAAGTTGTTA 5503

Qy        344 LysLysGlnGlnHis---SerLysLysArgThrAlaGlnAlaAspValSerAsp------ 360
Db       5504 CCTGAAAACAATTATGTCAGCCGGATATTGAGTACAATGGGTGCAACCTCTGAAAATGCT 5563

Qy        361 AlaLysLeuCys---------------------------------------------- 364
Db       5564 TCCAAAAGTTGTGACTCTGATCAAGGTAATAGTGAATCAACTGATAGTGGGTTTGACAGA 5623

Qy        365 --------------ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsn 379
Db       5624 ACTCCATTTAAGGGTAAGCAGAGAAACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCA 5683

Qy        380 AlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro 399
Db       5684 TCACTTCCTTGTGAAACTTCACAA------GAAGGTATCAAGGAGCATGATGCAGATCCT 5737

Qy        400 CysGluAspAspArgSerThr---------------------------------IlePro 408
Db       5738 AGTAAG------AGATCAACTCCTGCGCACTCTTTATTCACTGGAAACGATTCTGTTCCT 5791

Qy        409 ValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGly 428
Db       5792 TGTCCT--------------------CCGGGTACTCAGAGAACA------GAGAGGAAG 5824

Qy        429 LeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySer 448
Db       5825 CTCAGTTTACCCAAGAAGAAGACAAAGAAG---------CCTGTAATCGATAATGGGAAG 5875

Qy        449 SerLeuMet------------AsnTrpLeuAsnGlyLysLys------LysArgThrGly 462
Db       5876 AGCACTGTGATCAGTTTTAGTAACGGCATTGATGGAAGTCAAGTTAACTCGCATACTGGT 5935

Qy        463 SerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPro 482
Db       5936 CCTTCCATGAACACAGTATCCCAAACTCGAGACTTATTGAATGGGAAAAGGGTGGGCGGT 5995

Qy        483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAsp--------- 499
Db       5996 TTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGAAAATATCTCTCTCAGGTTAAT 6055

Qy        500 ------------ThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSer 515
Db       6056 GATAAGCCGATAACATCTTTGCATTTGCAAGACAATGATTATGTGAGGTCAAGAGACGCG 6115

Qy        516 ThrGlnArgCys-----------------SerSerLysGlyLysThrAlaGly---Leu 528
Db       6116 GAACCAAACTGTCTTCGAGATTTTAGTTCCTCTTCTAAATCCAGCTCAGGTGGATGGTTG 6175

Qy        529 SerLysGly---------------------------------------------LysThr 533
Db       6176 AGAACTGGAGTAGATATTGTTGACTTCAGAAACAACAACCACAATACAAACAGATCGTCT 6235

Qy        534 HisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHis 553
Db       6236 TTCTCGAACTTGAAGCTAAGATACCCCCCTTCTTCTACTGAAGTTGCGGATTTATCTCGG 6295

Qy        554 ValLeuSerAlaGluAspGln-CysGlnMetGluThrGluAsnSerValLeuSerHisSe 573
Db       6296 GTGCTGCAAAAGGTACAAGCTTTGCATCTATTTTTCTTCAATCGCCATGCTTTCTCATTC 6355

Qy        573 r------------------------------------------------------------ 573
Db       6356 TGTTTTCCAACTTGCACCTATTCTTGTATATTCCAAACTGGTGTGCTTCAGTGATTGGAT 6415

Qy        574 --------------------------Ala-LysValSerProAlaGluHisAspIleG 584
Db       6416 CAGGATGATAACGGGATTGATTTACTTGGCAGGATGCTTCTGGTGCAGATAGAAAGGGGA 6475

Qy        584 lnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuG 604
```



RESULT 3
US-09-774-528-286
; Sequence 286, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 286
; LENGTH: 6745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6147)
US-09-774-528-286

Alignment Scores:

Pred. No.:	4.41e-10	Length:	6745
Score:	219.00	Matches:	213
Percent Similarity:	33.7%	Conservative:	172
Best Local Similarity:	18.7%	Mismatches:	410
Query Match:	4.0%	Indels:	348
DB:	3	Gaps:	53

US-09-828-068-2 (1-1057) x US-09-774-528-286 (1-6745)

```
Qy        83 HisAspGln---------LysLysCysAspGluHisLysAlaSerSerSerProPheSer 99
             ||||||:::          ||||||   |||        :::
Db      3034 CATGATGAAGTGGTTGATGTCAAATGCCATTCACACTCTCCTTTTAGGAATGAAAGAGGA 3093

Qy       100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
               |||   |||                 |||::::::||||||      ::::::
Db      3094 AAAGGAAAAATAAGGCATCAT------ATATCCTGTATTGAAAAGTTAAGCAAAACAGAA 3147

Qy       120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
             :::  :::  |||                         |||||||||       |||
Db      3148 AGTATATCAGTACCC-------------------------ACCAGTGATCACAGGAGC 3180

Qy       140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer--- 158
             :::                    :::   |||:::      :::   |||||||||
Db      3181 CTC--------------------ATTGAAGCAAATCAAAGCAATTCCAAAGTTTCTGAA 3219

Qy       159 ----------------------ProSerThrGlnSerSer---------GlnGlyLys 167
                                   |||         ||||||           |||:::
Db      3220 CTTGACACAATTTATTGTACCTTGCCAAGAAAATCAAGCAGTTTTCTCATACATGGCAGG 3279

Qy       168 AsnAlaAspArgSerThrLeuProLysSerValGlnGluGly------------------ 181
               :::          :::      |||::::::   |||
Db      3280 CAGTCAGGAAGTAAAATAATGGCTGCTTCATTGAGGAATGGGCCACCTCCCTTCCAAATC 3339

Qy       182 -------------------------------------AsnAspSerLysCysAsnAla 188
                                                   |||          ::::::
Db      3340 AAAAATAATGTGGAAGATGCAATGGGGAACTATATGTTAAACAAATTTAGTCCCAGTTCT 3399

Qy       189 ProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln 208
             |||         |||      ::::::     :::||||||             |||:::
Db      3400 CCTGAGTCAGCGAATGAATGTTCCAAAGTCCTTTCAGACTCAGCC---------CTGGAA 3450

Qy       209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
               |||          :::      |||      :::||||||||||
Db      3451 GCACCTGAAGCCACAGAGAGAATGACAAATGTAAAAAGCAGTGGATCTACTTCCGTTAGA 3510

Qy       229 ValGlyAlaLeuPro--------------------GluValProGlnIleThrTrpHis 241
               |||   ||||||                    |||          |||
Db      3511 AAAGGACCACTTCCATTCCTCATCAACAGGGCTATGTCATGTCCCTCAGGGGAGCCACAT 3570

Qy       242 IleGlu-----------------------------ValAsnGlyAlaAspGlnProPro 251
                                               :::|||   |||
Db      3571 GCCTCAACTGGAAGAGAAGGAAGAAAAAAGCCATTGACCTCAGGCATGGATGCTTCTGAG 3630

Qy       252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
             ||||||:::       |||   ::::::          |||   ::::::
Db      3631 CTAACACCAAGGGCTTGGGAGAGAATCATTAGCCCTGTGGAAAGTGACTCATCT------ 3684

Qy       272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys--------------------- 284
                                  |||:::|||||||||
Db      3685 --------------GTTAGAGATTGTTCTTTAACCAAAAGACAACACCAAAAGGAAAAC 3729

Qy       284 ------------------------------------------------------------ 284

Db      3730 TTCCAAGAATACACTGAGAAAGAGGGTAAAATGGCTGCCTCCAGGAGAAGTGTATTTGCT 3789

Qy       285 ------------AspProAsnPro------------MetSerGlyLysGluArgAspGln 296
                         ||||||   |||            :::|||||||||||||||   :::
Db      3790 CTTTCAAATGAAGACCCTTTACCTTTTTGCTCAGACTTGTCAGGAAAAGAACGTGGGAAA 3849

Qy       297 ValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGlu 316
                    :::          |||                |||||||||   :::
Db      3850 ACATTACATAAAGTTAAG---ACGACTAGTACGTTTTCTGTTTCTGGTGATGAAGATAAT 3906
```

```
Qy       317 GlnIleCysAsnGluPro-CysGluGluValValLeuLysArgSerSerLysSerLys-- 335
                ||| |||       :::  :::  |||  |||  :::  :::|||
Db      3907 GTAAAATGTCTGGAGGTGGTCTCAATATATTACACTCTACCGAGGAAACCCAGCAAAAAA 3966

Qy       335 ------------------------------------------------------------ 335

Db      3967 TTCTGTAACCTCCTTCAACAGTATACACAAAATACTAATTTACTTATAGAATCACCTCAA 4026

Qy       336 ------------------------------ArgLysThrAspLysLysLeuMet----- 343
                                            :::||||||::::::   ::::::
Db      4027 GTGGAGACTGAAACATTTCCTAACGCTTTAGAAAAAGACAAACAGAATTATTCTACACGA 4086

Qy       344 -LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLe 363
              |||:::|||         :::      :::                 :::   ||
Db      4087 GAGCAGTCAGGAACACCTTCATGTGAAAATCTAAAGATGT---------CCGTCAACTCT 4137

Qy       363 uCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnVa 383
             |  ||||||  |||  |||:::  |||  |||:::      |||       ::
Db      4138 GATCAGACGCTCACCACTGAAAATATGACTGCCTTCCGATTATCA-AATAGGGGGCCCCT 4196

Qy       383 lGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAs 403
             :      :::      :::              |||   |||
Db      4197 AGCTCCTACATTACAGGAAATGGCTTCTGTTGAGGCAGCT-------------------- 4236

Qy       403 pArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHisTh 423
                     :::   :::|||   |||   |||             :::   :::
Db      4237 ----------GTTTCTCTTCCTGAAGAGGAATCTAAAGCTAGAGAGATTTTTTCAGATAA 4286

Qy       423 rValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspVa 443
              :::   :::       |||      |||:::||||||   :::|||
Db      4287 TTTAGCTAAAACACCTCTAGGTGATTCAGAAAACAAGAAGGAAAGA-------------- 4332

Qy       443 lValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySe 463
                                                 |||||||||   ::::::   ::
Db      4333 ----------------------------------GGCAAAAAGTTGCAAAGTGAAAC 4355

Qy       463 rValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProTh 483
             ::::|||   ::::::                |||:::|||   :::|||   :::
Db      4356 CCTGCATACTTCATTGATGCTTCAGAGAAAAAATGTATCCGAAGAAAAA---TCTGAAAA 4412

Qy       483 rAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHi 503
                          :::      |||      :::      |||      :::
Db      4413 TTGTCAACAATCCATTAATTCAAGTAACAGTGGTCCCTCTAGTCTTCCAGCT---CTTTC 4469

Qy       503 sLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGl 523
             :::  ::::::          :::      |||:::      |||:::      ||
Db      4470 AGAAGTTAATATTGGAAATTCCCAAACTAGAAGAAGTTCTTGGGAGTGTACAGGGAGTGG 4529

Qy       523 yLysThr---------------AlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSe 538
             |:::                 :::|||            |||   |||::::::   ::
Db      4530 TAGAGCCATTCCATTTACTGGAAGTGGCAAGTGTCCCCAGAAAGATCACACATCCACAGC 4589

Qy       538 rThrLysTyrGly------GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSe 556
             :         |||      |||      |||   |||:::
Db      4590 TGTAGGTGATGGCTCCAGTGGATCACAGCCTAGGGAAGGCAGAGGG------------GA 4637

Qy       556 rAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys-- 575
                           ||||||      |||   |||      ||||||||||||
Db      4638 CATTGGAACCAACTGCCAAAAAATGACT---AATAAAACACTTTCTCACTCAGAGAGTCA 4694

Qy       576 ----------ValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGlu-- 591
                     ::::::||||||   |||  :::|||:::      :::
Db      4695 AGTCTTTGCCCTTACTCCAGCATTGCATAAACTACAGCTTGGTGAGGAGACTCAGTCAGA 4754

Qy       592 ----------------------GlnSerLeuProLysLys------LysLysLysGlnLy 602
                                   ||||||   |||::::::      ::::::   |||::
Db      4755 TGAACCAAACTTAGAGAGTCTGCAGTCTGAACCAAGAGAATTACCTCAAAGAAGTCAGGA 4814

Qy       602 sLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGl 622
             :      :::|||      ::::::
Db      4815 GGCAAATATGACAGAGAGCAGGAAG---------------------------------- 4839
```

```
Qy       622 uLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIl 642
Db      4840 -------GCTGAAGATGAAATGCAGAAGTCAGCTTGGGATCAACCTTCACTTCCTGAAGG 4892

Qy       642 eAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAs 662
Db      4893 AAACAAA------AATAAAACCAACTTGGATGAC---------CTAGTAAAGGGGGAAAA 4937

Qy       662 pGlySerAspTyr--------------------------AlaSerSerValPheAspTh 673
Db      4938 TAGATCTTCAGTTAAACACAGATTGGCAGCCATGTCTAAAGCCAGCAGAAAATTCCCAGC 4997

Qy       673 rAsn------SerGlnGlnLysSerLeuAlaSer---------GlnSerThrGlnLysGl 688
Db      4998 TAAAGATGTAAGCCCCAGAAGACATGTAGCTACTATCTTCCCCCAAAGTGGAAGCAGATC 5057

Qy       688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGlu-----------------SerProHi 702
Db      5058 TGGCTTTGACCATTTATCTCTTGGCACAGTGGAGTGCAACCCACTGTTCCCTGAGCCTAC 5117

Qy       702 sProGlnAsnPheGlnSerThrGlnGluGlnGln---------------ThrHisLeuAr 717
Db      5118 TCCAAAATCTGCAGAGTCCATTGGCGAAAGCAGGTTGAGTGAGAACGGAAAGCATGTGAA 5177

Qy       717 gMetGluGlu-------------------------------------------MetVa 722
Db      5178 GAAATCCGAGAACCTTCTCCCCATTACTGTACTACCCAACAGAGAACCTTCTACACACGT 5237

Qy       722 lThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyr------IleAl 740
Db      5238 CAGCAACCAGAAGTCTAACAGCATTTCACAACGACATCAGAATGAGTTTAAAAACGTCTC 5297

Qy       740 aGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPh 760
Db      5298 AGAATCACCATCAAAGCATGAGAATTCTAAAGAC------------------------- 5331

Qy       760 eLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIl 780
Db      5332 ----GTCACAGCAGCTCAGAATTTAGTAAGAGAATCAGGAGCCCCA------------- 5373

Qy       780 eGlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerAr 800
Db      5374 -TCACCCATCACATTCACCAGCCTCAGGGAAGCAGAATTCTCTGACAATCAGAGGAGGCT 5432

Qy       800 gGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAl 820
Db      5433 GAGCCCTCCTTTT---CCACTGGAGCCT---GCACAGAAA------------------- 5466

Qy       820 aArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAs 840
Db      5467 ------------TCTAGAGTAAGCAGTCCACTGGCCAGTTTTCTGCAGCAACAAAGGAG 5513

Qy       840 nAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMe 860
Db      5514 TGCTTCATCTCTG--------------GAGTGGGAACCTGAGCCACACCTCTATCGT-- 5556

Qy       860 tMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMe 880
Db      5557 -------TCAAAGAGTTTAAAAAGCATTAATGTTCATGGCGATCTACTACGAAAAAGC-- 5607

Qy       880 tGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerTh 900
Db      5608 ------------CATCCTCCAAAAGTCAGGGAGCGCCATTTTTCTGAAAGCACTTCTAT 5654

Qy       900 rSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHi 920
Db      5655 T---------GACAATGCCCTGAGTCGACTGACCCTTGGGAATGAATTCTCTGTCAACAA 5705

Qy       920 sGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySe 940
Db      5706 TGGGTACAGTCGAAGATTCAGATCTTTTTCTGAACTCCCCTCCTGTGAT------GGAAA 5759

Qy       940 rLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLy 960
Db      5760 T-----------------GAAAGTTGGGCTTATCGCAGCGGGACAAAAACAGGTCCCAG 5801

Qy       960 sLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAl 980
Db      5802 GTCTGCAATATCTATA---------------------------------TACAGACC 5825

Qy       980 aLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSe 1000
Db      5826 TATCGACTATGGGATCTTTGGGAAAGAACAACAGTTAGCTTTCTTAGAAAATGTAAAGAG 5885

Qy      1000 rSerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGl 1020
Db      5886 GTCA---------CTTACACAAGGAAGATTATGGAAACCAAGTTTTCTTAAGAACCCTGG 5936

Qy      1020 yLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAs 1040
Db      5937 C-----------------------TTCCTAAAAGATGATTTG-----------AGGAA 5960

Qy      1040 nPro 1041
Db      5961 CCCT 5964
```

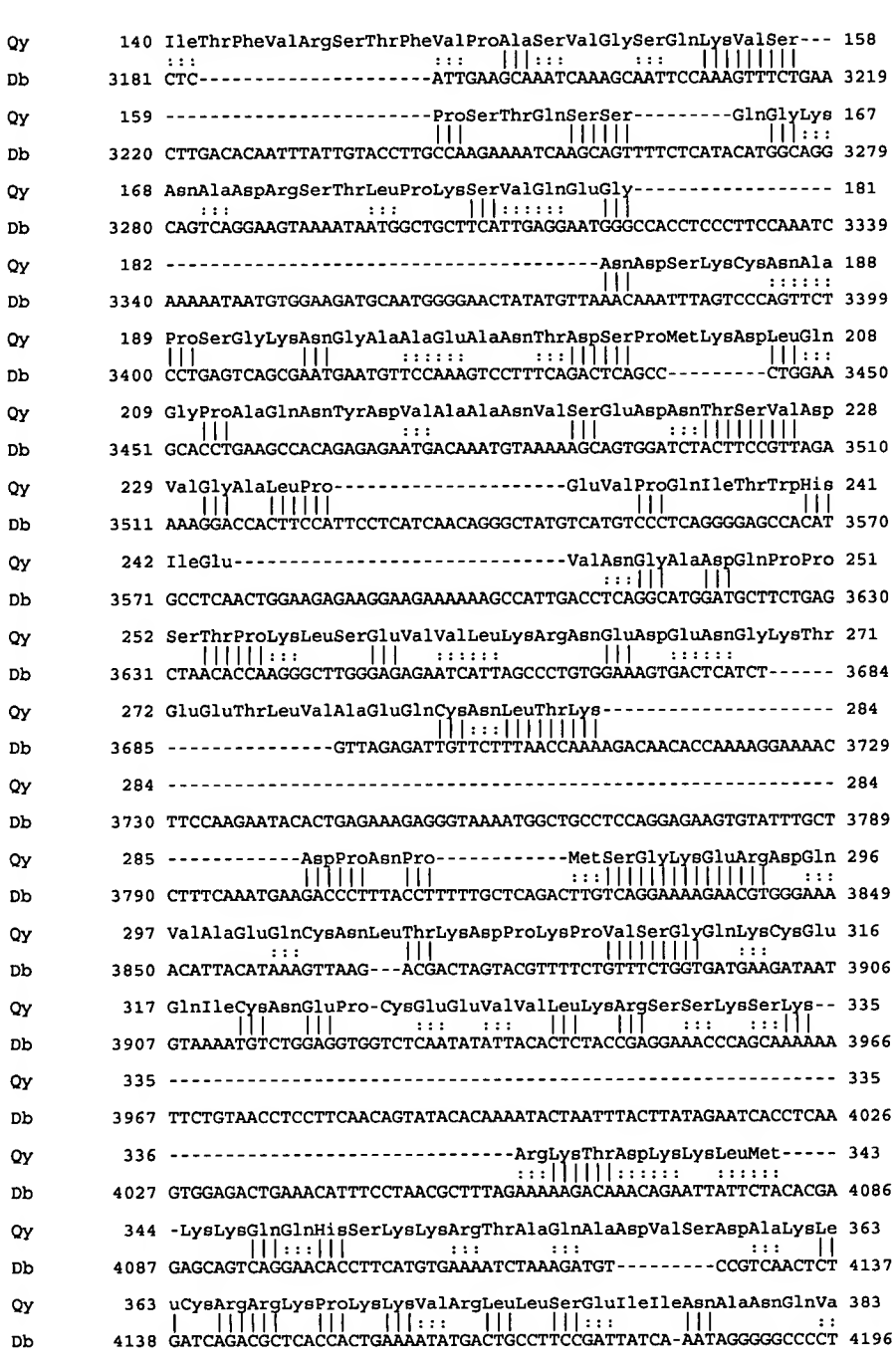
RESULT 4
US-10-120-988-286
; Sequence 286, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6919193el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 286
; LENGTH: 6745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6147)
US-10-120-988-286

Alignment Scores:

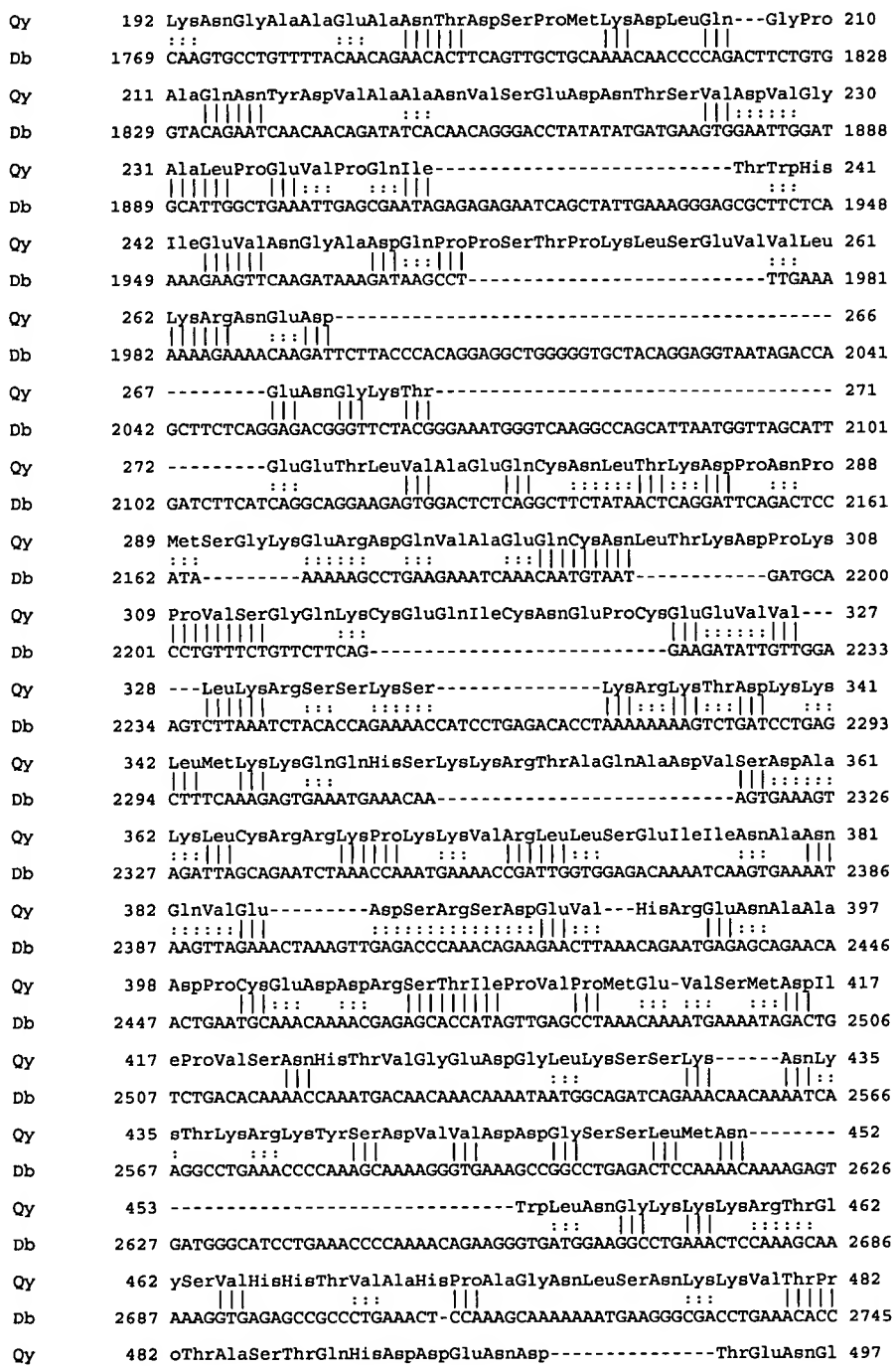
Pred. No.:	4.41e-10	Length:	6745
Score:	219.00	Matches:	213
Percent Similarity:	33.7%	Conservative:	172
Best Local Similarity:	18.7%	Mismatches:	410
Query Match:	4.0%	Indels:	348
DB:	3	Gaps:	53

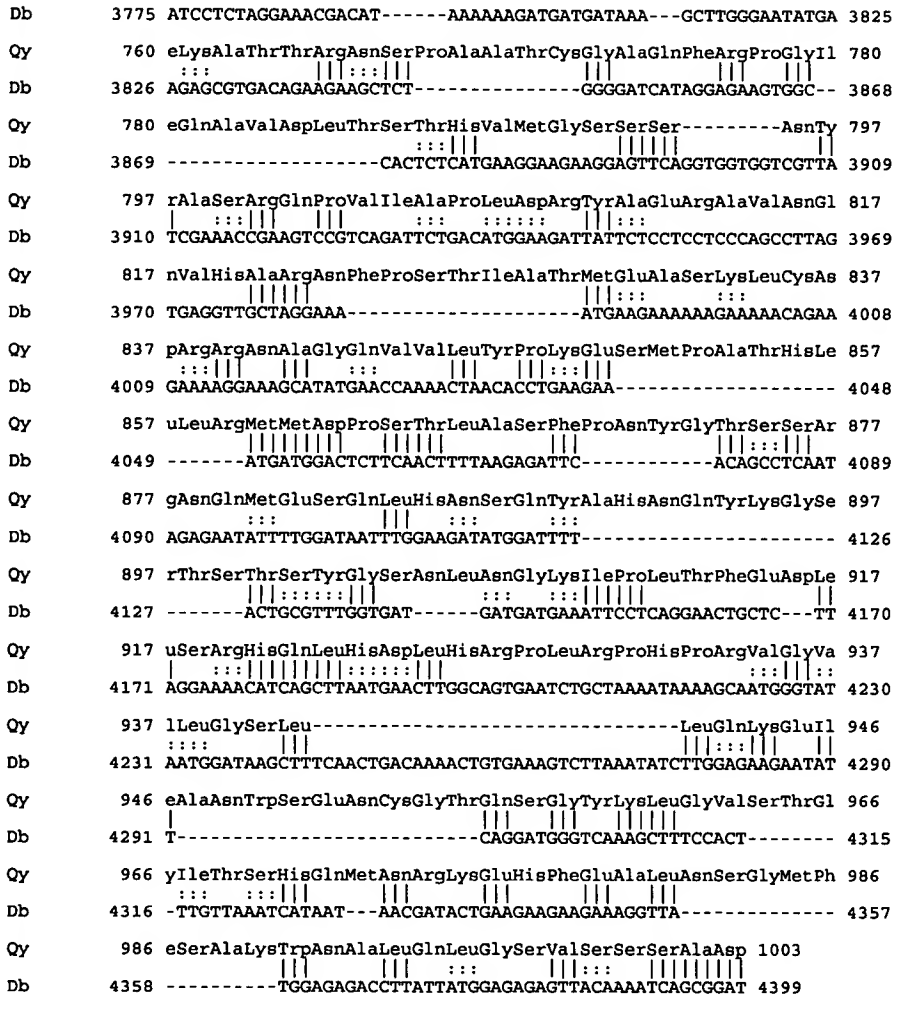
US-09-828-068-2 (1-1057) x US-10-120-988-286 (1-6745)

```
Qy        83 HisAspGln---------LysLysCysAspGluHisLysAlaSerSerSerProPheSer 99
Db      3034 CATGATGAAGTGGTTGATGTCAAATGCCATTCACACTCTCCTTTTAGGAATGAAAGAGGA 3093

Qy       100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
Db      3094 AAAGGAAAAATAAGGCATCAT------ATATCCTGTATTGAAAAGTTAAGCAAAACAGAA 3147

Qy       120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
Db      3148 AGTATATCAGTACCC-------------------------ACCAGTGATCACAGGAGC 3180
```


Qy 140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer--- 158
Db 3181 CTC--------ATTGAAGCAAATCAAAGCAATTCCAAAGTTTCTGAA 3219
Qy 159 ----------ProSerThrGlnSerSer---------GlnGlyLys 167
Db 3220 CTTGACACAATTTATTGTACCTTGCCAAGAAAATCAAGCAGTTTTCTCATACATGGCAGG 3279
Qy 168 AsnAlaAspArgSerThrLeuProLysSerValGlnGluGly---------- 181
Db 3280 CAGTCAGGAAGTAAAATAATGGCTGCTTCATTGAGGAATGGGCCACCTCCCTTCCAAATC 3339
Qy 182 ----------AsnAspSerLysCysAsnAla 188
Db 3340 AAAAATAATGTGGAAGATGCAATGGGGAACTATATGTTAAACAAATTTAGTCCCAGTTCT 3399
Qy 189 ProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln 208
Db 3400 CCTGAGTCAGCGAATGAATGTTCCAAAGTCCTTTCAGACTCAGCC---------CTGGAA 3450
Qy 209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
Db 3451 GCACCTGAAGCCACAGAGAGAATGACAAATGTAAAAAGCAGTGGATCTACTTCCGTTAGA 3510
Qy 229 ValGlyAlaLeuPro----------GluValProGlnIleThrTrpHis 241
Db 3511 AAAGGACCACTTCCATTCCTCATCAACAGGGCTATGTCATGTCCCTCAGGGGAGCCACAT 3570
Qy 242 IleGlu----------ValAsnGlyAlaAspGlnProPro 251
Db 3571 GCCTCAACTGGAAGAGAAGGAAGAAAAAAGCCATTGACCTCAGGCATGGATGCTTCTGAG 3630
Qy 252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
Db 3631 CTAACACCAAGGGCTTGGGAGAGAATCATTAGCCCTGTGGAAAGTGACTCATCT------ 3684
Qy 272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys---------- 284
Db 3685 --------------GTTAGAGATTGTTCTTTAACCAAAAGACAACACCAAAAGGAAAAC 3729
Qy 284 ---------- 284
Db 3730 TTCCAAGAATACACTGAGAAAGAGGGTAAAATGGCTGCCTCCAGGAGAAGTGTATTTGCT 3789
Qy 285 ----------AspProAsnPro----------MetSerGlyLysGluArgAspGln 296
Db 3790 CTTTCAAATGAAGACCCTTTACCTTTTGCTCAGACTTGTCAGGAAAAGAACGTGGGAAA 3849
Qy 297 ValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGlu 316
Db 3850 ACATTACATAAAGTTAAG---ACGACTAGTACGTTTTCTGTTTCTGGTGATGAAGATAAT 3906
Qy 317 GlnIleCysAsnGluPro-CysGluGluValValLeuLysArgSerSerLysSerLys-- 335
Db 3907 GTAAAATGTCTGGAGGTGGTCTCAATATATTACACTCTACCGAGGAAACCCAGCAAAAAA 3966
Qy 335 ---------- 335
Db 3967 TTCTGTAACCTCCTTCAACAGTATACACAAAATACTAATTTACTTATAGAATCACCTCAA 4026
Qy 336 ----------ArgLysThrAspLysLysLeuMet----- 343
Db 4027 GTGGAGACTGAAACATTTCCTAACGCTTTAGAAAAAGACAAACAGAATTATTCTACACGA 4086
Qy 344 -LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLe 363
Db 4087 GAGCAGTCAGGAACACCTTCATGTGAAAATCTAAAGATGT---------CCGTCAACTCT 4137
Qy 363 uCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnVa 383
Db 4138 GATCAGACGCTCACCACTGAAAATATGACTGCCTTCCGATTATCA-AATAGGGGGCCCCT 4196
Qy 383 lGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAs 403
Db 4197 AGCTCCTACATTACAGGAAATGGCTTCTGTTGAGGCAGCT---------- 4236
Qy 403 pArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHisTh 423
Db 4237 ----------GTTTCTCTTCCTGAAGAGGAATCTAAAGCTAGAGAGATTTTTTCAGATAA 4286
Qy 423 rValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspVa 443
Db 4287 TTTAGCTAAAACACCTCTAGGTGATTCAGAAAACAAGAAGGAAAGA---------- 4332
Qy 443 lValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySe 463
Db 4333 ----------GGCAAAAAGTTGCAAAGTGAAAC 4355
Qy 463 rValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProTh 483
Db 4356 CCTGCATACTTCATTGATGCTTCAGAGAAAAAATGTATCCGAAGAAAAA---TCTGAAAA 4412
Qy 483 rAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHi 503
Db 4413 TTGTCAACAATCCATTAATTCAAGTAACAGTGGTCCCTCTAGTCTTCCAGCT---CTTTC 4469
Qy 503 sLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGl 523
Db 4470 AGAAGTTAATATTGGAAATTCCCAAACTAGAAGAAGTTCTTGGGAGTGTACAGGGAGTGG 4529
Qy 523 yLysThr----------AlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSe 538
Db 4530 TAGAGCCATTCCATTTACTGGAAGTGGCAAGTGTCCCCAGAAAGATCACACATCCACAGC 4589
Qy 538 rThrLysTyrGly------GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSe 556
Db 4590 TGTAGGTGATGGCTCCAGTGGATCACAGCCTAGGGAAGGCAGAGGG----------GA 4637
Qy 556 rAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys-- 575
Db 4638 CATTGGAACCAACTGCCAAAAAATGACT---AATAAAACACTTTCTCACTCAGAGAGTCA 4694
Qy 576 ----------ValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGlu-- 591
Db 4695 AGTCTTTGCCCTTACTCCAGCATTGCATAAACTACAGCTTGGTGAGGAGACTCAGTCAGA 4754
Qy 592 ----------GlnSerLeuProLysLys------LysLysLysGlnLy 602
Db 4755 TGAACCAAACTTAGAGAGTCTGCAGTCTGAACCAAGAGAATTACCTCAAAGAAGTCAGGA 4814
Qy 602 sLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGl 622
Db 4815 GGCAAATATGACAGAGAGCAGGAAG---------- 4839
Qy 622 uLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIl 642
Db 4840 -------GCTGAAGATGAAATGCAGAAGTCAGCTTGGGATCAACCTTCACTTCCTGAAGG 4892
Qy 642 eAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAs 662
Db 4893 AAACAAA------AATAAAACCAACTTGGATGAC---------CTAGTAAAGGGGGAAAA 4937
Qy 662 pGlySerAspTyr----------AlaSerSerValPheAspTh 673
Db 4938 TAGATCTTCAGTTAAACACAGATTGGCAGCCATGTCTAAAGCCAGCAGAAAATTCCCAGC 4997
Qy 673 rAsn------SerGlnGlnLysSerLeuAlaSer---------GlnSerThrGlnLysGl 688
Db 4998 TAAAGATGTAAGCCCCAGAAGACATGTAGCTACTATCTTCCCCCAAAGTGGAAGCAGATC 5057
Qy 688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGlu----------SerProHi 702
Db 5058 TGGCTTTGACCATTTATCTCTTGGCACAGTGGAGTGCAACCCACTGTTCCCTGAGCCTAC 5117
Qy 702 sProGlnAsnPheGlnSerThrGlnGluGlnGln----------ThrHisLeuAr 717

```
Db       5118 TCCAAAATCTGCAGAGTCCATTGGCGAAAGCAGGTTGAGTGAGAACGGAAAGCATGTGAA 5177

Qy        717 gMetGluGlu----------------------------------------------MetVa 722

Db       5178 GAAATCCGAGAACCTTCTCCCCATTACTGTACTACCCAACAGAGAACCTTCTACACACGT 5237

Qy        722 lThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyr------IleAl 740

Db       5238 CAGCAACCAGAAGTCTAACAGCATTTCACAACGACATCAGAATGAGTTTAAAAACGTCTC 5297

Qy        740 aGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPh 760

Db       5298 AGAATCACCATCAAAGCATGAGAATTCTAAAGAC------------------------- 5331

Qy        760 eLysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIl 780

Db       5332 ----GTCACAGCAGCTCAGAATTTAGTAAGAGAATCAGGAGCCCCA-------------- 5373

Qy        780 eGlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerAr 800

Db       5374 -TCACCCATCACATTCACCAGCCTCAGGGAAGCAGAATTCTCTGACAATCAGAGGAGGCT 5432

Qy        800 gGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAl 820

Db       5433 GAGCCCTCCTTTT---CCACTGGAGCCT---GCACAGAAA-------------------- 5466

Qy        820 aArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAs 840

Db       5467 ------------TCTAGAGTAAGCAGTCCACTGGCCAGTTTTCTGCAGCAACAAAGGAG 5513

Qy        840 nAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMe 860

Db       5514 TGCTTCATCTCTG--------------GAGTGGGAACCTGAGCCACACCTCTATCGT-- 5556

Qy        860 tMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMe 880

Db       5557 -------TCAAAGAGTTTAAAAAGCATTAATGTTCATGGCGATCTACTACGAAAAAGC-- 5607

Qy        880 tGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerTh 900

Db       5608 ------------CATCCTCCAAAAGTCAGGGAGCGCCATTTTTCTGAAAGCACTTCTAT 5654

Qy        900 rSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHi 920

Db       5655 T---------GACAATGCCCTGAGTCGACTGACCCTTGGGAATGAATTCTCTGTCAACAA 5705

Qy        920 sGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySe 940

Db       5706 TGGGTACAGTCGAAGATTCAGATCTTTTTCTGAACTCCCCTCCTGTGAT------GGAAA 5759

Qy        940 rLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLy 960

Db       5760 T-----------------GAAAGTTGGGCTTATCGCAGCGGGACAAAAACAGGTCCCAG 5801

Qy        960 sLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAl 980

Db       5802 GTCTGCAATATCTATA---------------------------------TACAGACC 5825

Qy        980 aLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSe 1000

Db       5826 TATCGACTATGGGATCTTTGGGAAAGAACAACAGTTAGCTTTCTTAGAAAATGTAAAGAG 5885

Qy       1000 rSerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGl 1020

Db       5886 GTCA---------CTTACACAAGGAAGATTATGGAAACCAAGTTTTCTTAAGAACCCTGG 5936

Qy       1020 yLysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAs 1040

Db       5937 C----------------------TTCCTAAAAGATGATTTG-----------AGGAA 5960

Qy       1040 nPro 1041

Db       5961 CCCT 5964

RESULT 5
US-10-144-198-11
; Sequence 11, Application US/10144198
; Patent No. 6833247
; GENERAL INFORMATION:
;  APPLICANT: OriGene Technologies Inc
;  TITLE OF INVENTION: Regulated Prostate Cance Genes
;  FILE REFERENCE: 9U 105 R1
;  CURRENT APPLICATION NUMBER: US/10/144,198
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 8733
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (473)..(8566)
US-10-144-198-11

Alignment Scores:
Pred. No.:              2.71e-09        Length:         8733
Score:                  213.00          Matches:        233
Percent Similarity:     33.2%           Conservative:   165
Best Local Similarity:  19.4%           Mismatches:     411
Query Match:            3.9%            Indels:         391
DB:                     3               Gaps:           56

US-09-828-068-2 (1-1057) x US-10-144-198-11 (1-8733)

Qy          29 AlaProValLeuGluLeuThrAlaThrProArg-------------------------- 39

Db        1308 GCTCCCCTGCTGGAAGTGAA-GGAACTCCTAAAGGCTCAAGACCACCTTTAATCCTACAA 1366

Qy          40 --------------------------GlnAspAlaAlaAlaGluAlaGlyValAspGlu 50

Db        1367 TCTCAGTCTCTACCTTGTTCATCACCTCGAGATGTTCCACCAGATATCTTGCTAGATTCT 1426

Qy          51 ProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLys 70

Db        1427 CCAGAAAGAAAA-------------------------------------------CAAAAG 1444

Qy          71 LysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLysCysAspGlu 90

Db        1445 AAGCAGAAGAAAATG------------AAATTAGGCAAGGATGAAAAAGAGCAGAGTGAG 1492

Qy          91 HisLysAlaSer------------SerSerProPheSerValAlaLysPheArgArgTrp 106

Db        1493 AAAGCGGCAATGTATGATATAATTAGTTCTCCATCCAAG--------------------- 1531

Qy         107 AspCysSerLysCys---------LeuAspLysLeuLysThrSerAspAsnGlyThrAla 123

Db        1532 GACTCTACTAAACTTACATTAAGACTTTCTCGTGTAAGGTCTTCAGACATGGAC------ 1585

Qy         124 ProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheVal 143

Db        1586 -----------------CAGCAAGAGGATATGATTTCTGGTGTGGAAAATAGCAATGTT 1627

Qy         144 ArgSerThrPheValProAlaSerVal---------GlySerGlnLysValSerProSer 160

Db        1628 TCAGAAAATGATATTCCTTTTAATGTGCAGTACCCAGGACAGACTTCAAAAACACCCATT 1687

Qy         161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180

Db        1688 ACT-------------------------------------CCACAAGATATAAACCGC 1708

Qy         181 Gly---AsnAspSerLysCysAsnAla-----------------------ProSerGly 191

Db        1709 CCACTAAATGCTGCTCAATGTTTGTCGCAGCAAGAACAAACAGCATTCCTTCCAGCAAAT 1768
```


Qy 192 LysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln---GlyPro 210
Db 1769 CAAGTGCCTGTTTTACAACAGAACACTTCAGTTGCTGCAAAACAACCCCAGACTTCTGTG 1828
Qy 211 AlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGly 230
Db 1829 GTACAGAATCAACAACAGATATCACAACAGGGACCTATATATGATGAAGTGGAATTGGAT 1888
Qy 231 AlaLeuProGluValProGlnIle----------ThrTrpHis 241
Db 1889 GCATTGGCTGAAATTGAGCGAATAGAGAGAGAATCAGCTATTGAAAGGGAGCGCTTCTCA 1948
Qy 242 IleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeu 261
Db 1949 AAAGAAGTTCAAGATAAAGATAAGCCT----------TTGAAA 1981
Qy 262 LysArgAsnGluAsp---------- 266
Db 1982 AAAAGAAAACAAGATTCTTACCCACAGGAGGCTGGGGGTGCTACAGGAGGTAATAGACCA 2041
Qy 267 ---------GluAsnGlyLysThr---------- 271
Db 2042 GCTTCTCAGGAGACGGGTTCTACGGGAAATGGGTCAAGGCCAGCATTAATGGTTAGCATT 2101
Qy 272 ---------GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
Db 2102 GATCTTCATCAGGCAGGAAGAGTGGACTCTCAGGCTTCTATAACTCAGGATTCAGACTCC 2161
Qy 289 MetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys 308
Db 2162 ATA---------AAAAAGCCTGAAGAAATCAAACAATGTAAT----------GATGCA 2200
Qy 309 ProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValVal--- 327
Db 2201 CCTGTTTCTGTTCTTCAG----------GAAGATATTGTTGGA 2233
Qy 328 ---LeuLysArgSerSerLysSer----------LysArgLysThrAspLysLys 341
Db 2234 AGTCTTAAATCTACACCAGAAAACCATCCTGAGACACCTAAAAAAAAGTCTGATCCTGAG 2293
Qy 342 LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla 361
Db 2294 CTTTCAAAGAGTGAAATGAAACAA----------AGTGAAAGT 2326
Qy 362 LysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsn 381
Db 2327 AGATTAGCAGAATCTAAACCAAATGAAAACCGATTGGTGGAGACAAAATCAAGTGAAAAT 2386
Qy 382 GlnValGlu---------AspSerArgSerAspGluVal---HisArgGluAsnAlaAla 397
Db 2387 AAGTTAGAAACTAAAGTTGAGACCCAAACAGAAGAACTTAAACAGAATGAGAGCAGAACA 2446
Qy 398 AspProCysGluAspAspArgSerThrIleProValProMetGlu-ValSerMetAspIl 417
Db 2447 ACTGAATGCAAACAAAACGAGAGCACCATAGTTGAGCCTAAACAAAATGAAAATAGACTG 2506
Qy 417 eProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLys------AsnLy 435
Db 2507 TCTGACACAAAACCAAATGACAACAAACAAAATAATGGCAGATCAGAAACAACAAAATCA 2566
Qy 435 sThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetAsn-------- 452
Db 2567 AGGCCTGAAACCCCAAAGCAAAAGGGTGAAAGCCGGCCTGAGACTCCAAAACAAAAGAGT 2626
Qy 453 ----------TrpLeuAsnGlyLysLysLysArgThrGl 462
Db 2627 GATGGGCATCCTGAAACCCCAAAACAGAAGGGTGATGGAAGGCCTGAAACTCCAAAGCAA 2686
Qy 462 ySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrPr 482
Db 2687 AAAGGTGAGAGCCGCCCTGAAACT-CCAAAGCAAAAAAATGAAGGGCGACCTGAAACACC 2745
Qy 482 oThrAlaSerThrGlnHisAspAspGluAsnAsp----------ThrGluAsnGl 497
Db 2746 A------AAACACAGGCATGACAATAGGAGGGATTCTGGAAAGCCATCTACAGAGAAAAA 2799
Qy 497 yLeuAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGl 517
Db 2800 ACCTGAAGTGTCTAAACATAAACAAGATACTAAATCTGACTCACCTCGGTTAAAATCAGA 2859
Qy 517 nArgCysSerSer----------LysGlyLysThrAlaGlyLeuSerLysGlyLy 532
Db 2860 ACGAGCTGAAGCCTTAAAGCAGAGACCTGATGGGCGATCTGTTTCTGAGTCACTAAGACG 2919
Qy 532 sThrHisSer----------AlaAlaSerThrLysTyrGlyGlyGl 544
Db 2920 TGACCATGATAATAAACAAAAATCAGATGACAGGGGTGAATCAGAGCGACATCGAGGGGA 2979
Qy 544 uSerThrArg----------AsnGlyGlnAsnIleHisValLe 555
Db 2980 TCAGTCTAGGGTTCGAAGACCAGAAACATTGAGATCCTCTAGTAGAAATGAACATGGCAT 3039
Qy 555 uSerAlaGlu---------- 558
Db 3040 TAAATCTGATAGTTCAAAAACTGATAAACTAGAACGAAAACACAGGCATGAATCAGGGGA 3099
Qy 559 ----------AspGlnCysGlnMetGl 564
Db 3100 CTCAAGGGAAAGACCATCTTCTGGGGAACAAAAATCAAGACCTGACAGTCCTCGTGTTAA 3159
Qy 564 uThrGluAsnSerValLeuSerHisSerAlaLysVal---------SerProAlaGluHi 581
Db 3160 ACAAGGAGATTCTAATAAATCAAGATCTGATAAACTTGGTTTTAAATCACCAACTAGTAA 3219
Qy 581 sAspIleGlnIleMetSerAsp---------- 588
Db 3220 AGATGACAAAAGGACAGAGGGTAACAAGAGTAAAGTAGACACTAATAAAGCACACCCTGA 3279
Qy 589 ----------LeuHisGluGl 592
Db 3280 CAATAAGGCAGAATTTCCAAGTTATTTGTTGGGGGGCAGGTCTGGTGCGTTGAAAAATTT 3339
Qy 592 nSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThrMe 612
Db 3340 TGTCATTCCGAAAATCAAGAGGGATAAAGATGGCAATGTTACTCAGGAGACAAAGAAAAT 3399
Qy 612 t---IleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluAr 631
Db 3400 GGAAATGAAAGGAGAGCCGAAAGACAAAGTAGAAAAAATAGGATTA---------- 3445
Qy 631 gGlnLeuMetThrGluThrAspCysSerAspIleAsnArg----------Il 645
Db 3446 ----------GTTGAAGATCTAAATAAAGGAGCTAAGCCTGTAGTTGT 3483
Qy 645 eGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAs 665
Db 3484 GCTACAAAAACTGTCTTTGGATGAT------GTTCAGAAACTTATTAAAGATAGAGAGGA 3537
Qy 665 pTyrAlaSerSerValPhe----------AspThrAsnSerGlnGlnLysSerLe 680
Db 3538 CAAATCAAGAAGTTCCCTTAAACCTATCAAGAATAAACCATCAAAGTCAAATAAAGGTAG 3597
Qy 680 uAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSe 700
Db 3598 TATAGATCAATCAGTGTTAAAAGAATTACCCCCTGAACTC---CTGGCAGAAATTGAGTC 3654
Qy 700 rProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGl 720
Db 3655 CACCATGCCACTTTGTGAACGTGTGAAAATGAACAAACGCAAGCGTAGCACAGTTAATGA 3714
Qy 720 uMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAl 740
Db 3715 AAAGCCAAAATATGCTGAAATCAGTTCAGATGAAGATAATGATAGTGATGAAGCTTTTGA 3774
Qy 740 aGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPh 760

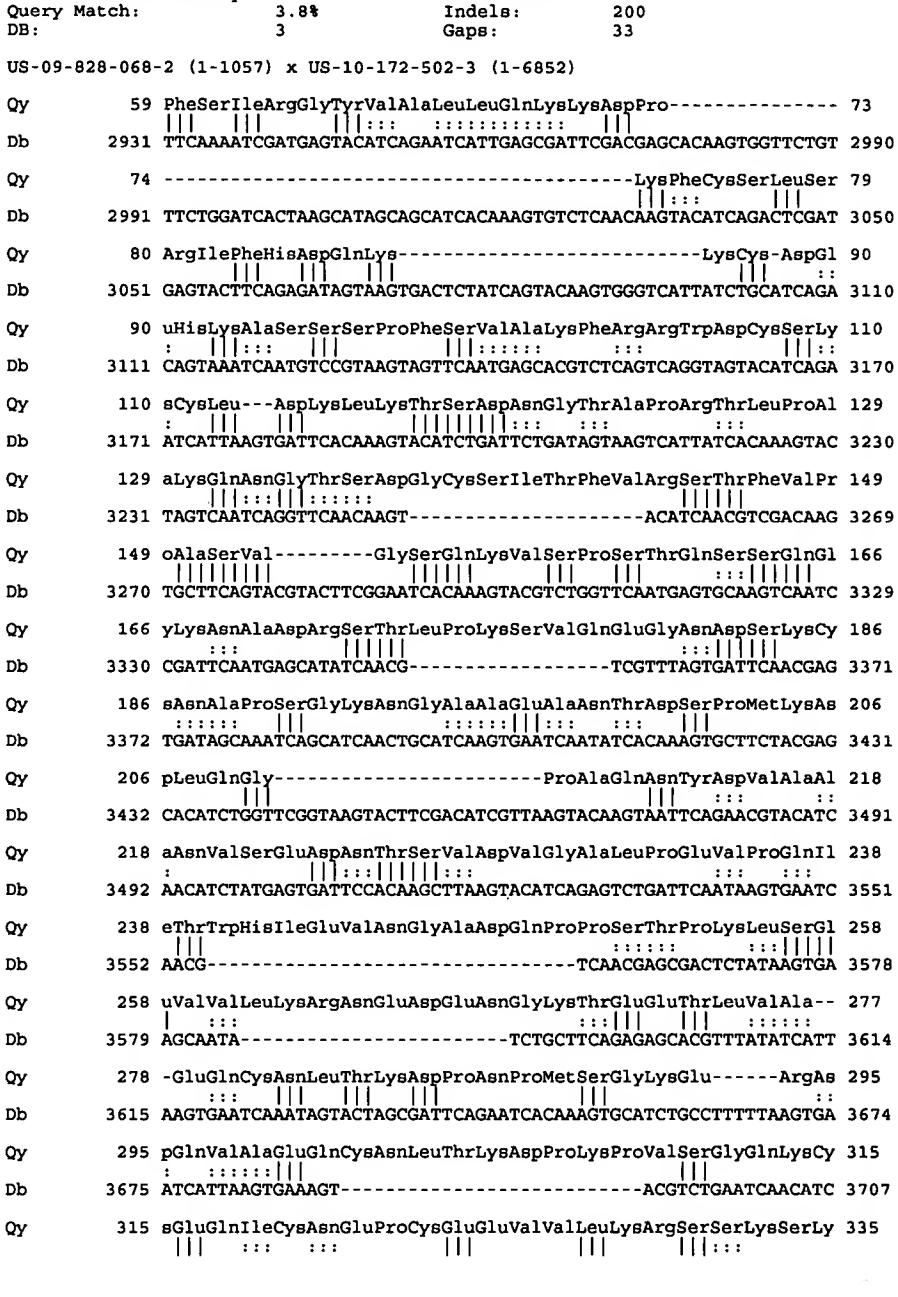


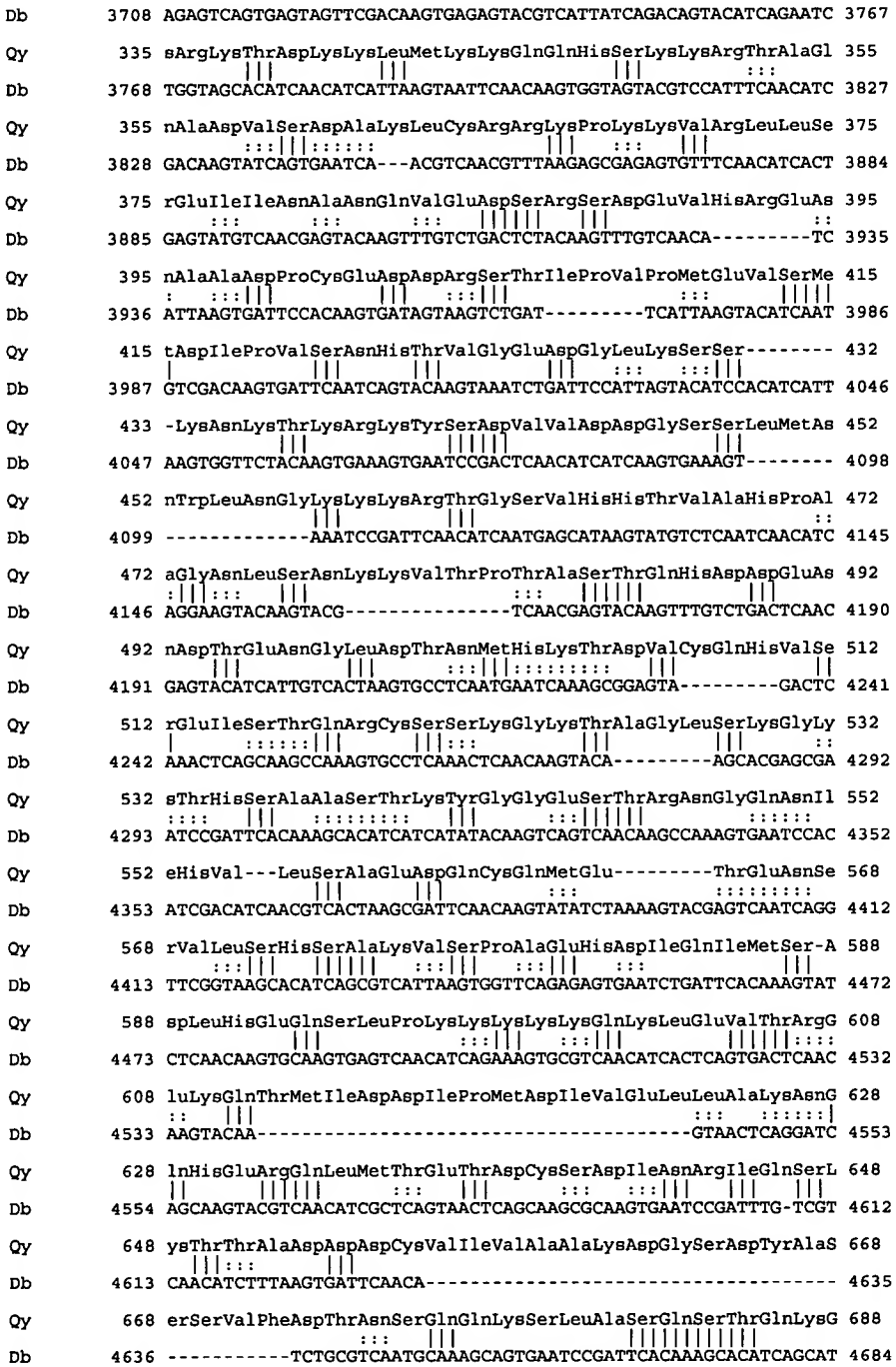
RESULT 6
US-10-172-502-3
; Sequence 3, Application US/10172502
; Patent No. 6841154
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 6852
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-3

Alignment Scores:

Pred. No.:	4.4e-09	Length:	6852
Score:	209.00	Matches:	192
Percent Similarity:	34.1%	Conservative:	152
Best Local Similarity:	19.0%	Mismatches:	467
Query Match:	3.8%	Indels:	200
DB:	3	Gaps:	33

US-09-828-068-2 (1-1057) x US-10-172-502-3 (1-6852)

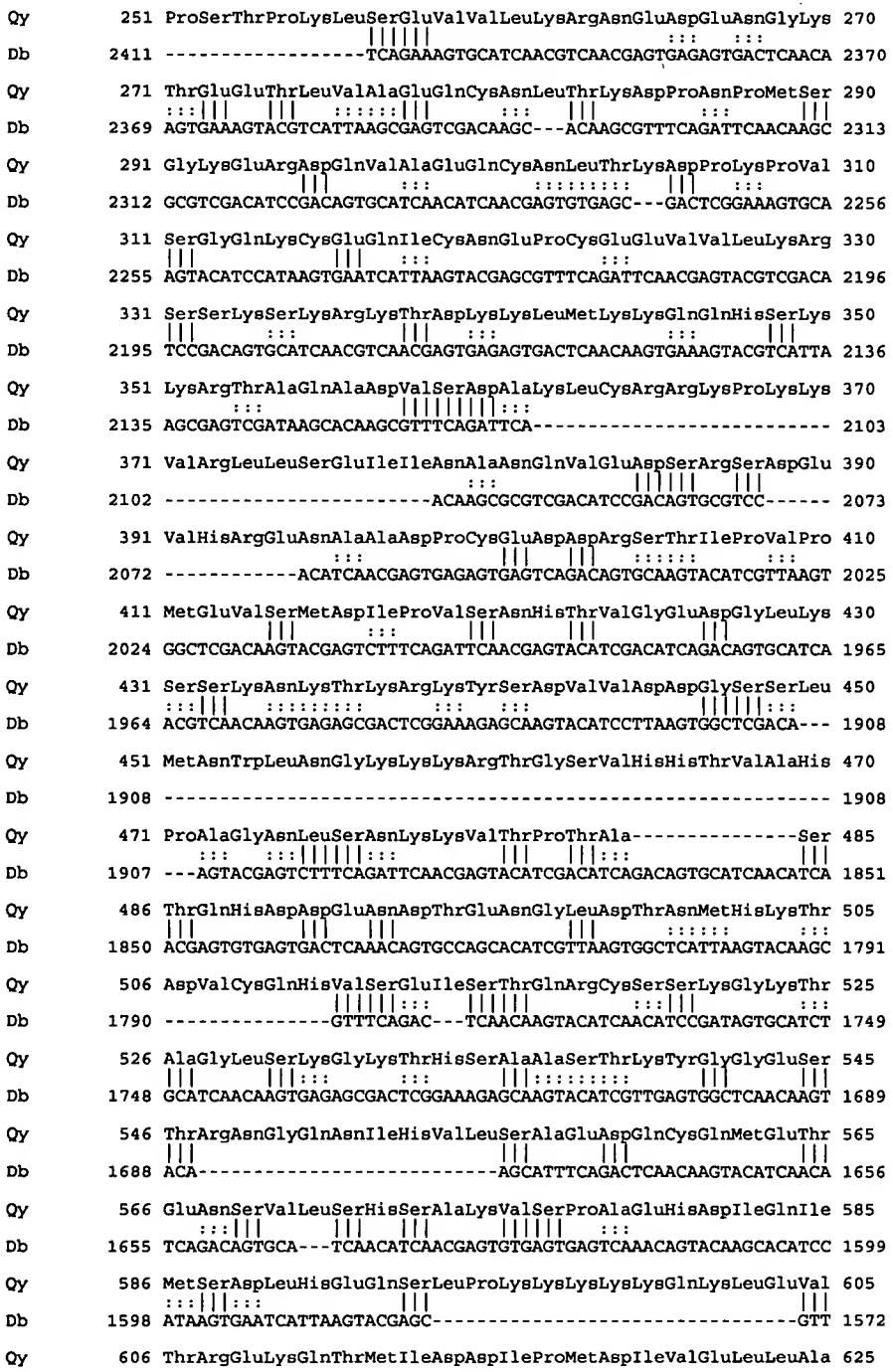




RESULT 7
US-09-134-001C-1626
; Sequence 1626, Application US/09134001C
; Patent No. 6380370

```
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1626
;   LENGTH: 6414
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1626

Alignment Scores:
Pred. No.:                3.4e-08        Length:          6414
Score:                    199.50         Matches:         186
Percent Similarity:       36.6%          Conservative:    185
Best Local Similarity:    18.3%          Mismatches:      503
Query Match:              3.6%           Indels:          146
DB:                       3              Gaps:            26

US-09-828-068-2 (1-1057) x US-09-134-001C-1626 (1-6414)

Qy        115 LeuLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThr 134
Db       1927 ATTACTACAACTGACAAT------AGTAAAGACGTATTAACTACAACTGTGACAGGTCTA 1980

Qy        135 SerAspGlyCysSerIleThrPheValArgSerThrPhe------ValProAlaSerVal 152
Db       1981 CCTTCAGGGTTATCTTTTGATCAAACGACAAATACAATTACTGGCACGCCAAGTGAAGTA 2040

Qy        153 GlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSer 172
Db       2041 GGAACTACAACTGTGACAGTTAATACTACTGATGCTACTGGGAACGTAACATCTAAGCAA 2100

Qy        173 ThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLys 192
Db       2101 TT-TAC------------------AATAACGATTCAAGATACAATCAGCCC--------- 2132

Qy        193 AsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGln 212
Db       2133 ------TGTTGTAAATGTGACGCCAAGTCAAGCATCAGAAGTTTTCACGCCGATTAATCC 2186

Qy        213 AsnTyrAsp-ValAlaAlaAsnValSer------------------------------- 221
Db       2187 AATTACGATAACTGCTACAGATAATAGTGGCAAAGTGGTAACGCATACAGTAACTGGATT 2246

Qy        222 ------------------GluAspAsnThrSerValAspValGlyAlaLeuProGluVa 235
Db       2247 GCCACAAGGACTTAAATTTGATGCATCTACGAATTCAATTGTTGGAACTCCAACTCAAAT 2306

Qy        235 lProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLy 255
Db       2307 AGGAACAAATACAATCACGATTGAGTCAACGGATGCGAGCGGAAATAAAACTACGACTAA 2366

Qy        255 sLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr----------- 271
Db       2367 AATTAATTATGAAGTAACGAGAAATAGCGCAAGTGACTCTACTTCCACTAGCATAGTAAA 2426

Qy        272 ----------------------GluGluThrLeuValAlaGluGlnCysAsnLeuThrLy 284
Db       2427 TAGTGTTTCAACAAGTATAAGTAATAGTACATCGCTAAGTGATAGTGTAAAAGCGAGTCA 2486

Qy        284 sAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuTh 304
Db       2487 ATCATTATCAACAAGTAAGAGTCTCTCAGAGTCGTTAAGTGCGTCGACGTCGAATAGTAC 2546

Qy        304 rLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCysGl 324

Db       2547 GTCAATCCAAGCGAGCGAATCAGCAAGCACAAGTAAACAATTA---TCAGAATCAGCGAG 2603

Qy        324 uGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMetLy 344
Db       2604 TACGTCGACATCTGATAGTGCATCAGAGTCAGCAAGAAAGAGTGAGTCA---------AC 2654

Qy        344 sLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAla-LysLeuC 364
Db       2655 GAGTAAAAGTACGTCATTAAGTGAATCGACAAGCACAAGCGTTTCAGATTCAGCAAGCGT 2714

Qy        364 ysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValG 384
Db       2715 GTCAACGT------CAGAAAGTGCGTCAACATCAACAAGTGTGAGTGGCTCAACAAGTAC 2768

Qy        384 luAspSerArgSerAspGluValHisArgGlu------AsnAlaAlaAspProCysGluA 402
Db       2769 AAGCATTTCAGACTCAACGAGCACATCGACG-TCTGATAGCGCATCAATAAAAGCGAGCG 2827

Qy        402 spAspArgSerThr---------IleProValProMetGluValSerMetAspIleProV 419
Db       2828 AATCAGCAAGCACAAGTAAACTATTATCAGAATCAGTGAGTACGTCGACGTCTGATAGCG 2887

Qy        419 alSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgL 439
Db       2888 CATCAACGTCAACGAGCGTAAGTGACTCCAATAGCGCAAGTACGTCATTAAGTAAATCGA 2947

Qy        439 ysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysL 459
Db       2948 CAAGCACAAGCGTTTCAGATTCAACAAGTACA---------------------------- 2979

Qy        459 ysArgThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysL 479
Db       2980 ----------------------TCGACATCCGACAGTGCGTCCACATCAACGAGTGAGA 3016

Qy        479 ysValThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuA 499
Db       3017 GTGAGTCAGACAGTGCAAGTACGTCATTAAGTGAATCGACAAGTACAAGCGTTTCAGATT 3076

Qy        499 spThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgC 519
Db       3077 CAACGAGTACGTCGACATCCGAC------------AGTGCATCAATGTCAGCGAGCGAGA 3124

Qy        519 ysSerSerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerT 539
Db       3125 GTGAATCAAACAGTAAAAGTACGTCATTAAGTGAA--------------TCGACAAGTA 3169

Qy        539 hrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluA 559
Db       3170 CGAGTCTTTCAGGC---TCAACGAGTGCGTCGACATCCGACAGTGCATCAACGTCAACGA 3226

Qy        559 spGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProA 579
Db       3227 GTGAGAGTGAGTCAGACAGTACAAGTACGTCATTAAGTGAATCGACAAGTACGAGTCTTT 3286

Qy        579 laGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysL 599
Db       3287 CAGGCTCAACGAGTGCGTCGACATCCGACAGTGCATCAACGTCAACAAGCGAGAGT---- 3342

Qy        599 ysLysGlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetA 619
Db       3343 -------------GACTCAACAAGTGAAAGTACATCGTTGAGTGAATCATTAAGCACAA 3388

Qy        619 spIleValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspC 639
Db       3389 GCGTTTCAGATTCAACAAGCGCGTCAACGTCAGAA---AGTGCATCCACATCAACGAGTG 3445

Qy        639 ysSerAspIleAsnArgIleGln--------------SerLysThrThrAlaAspAspA 654
Db       3446 AGAGTGAGTCAAACAGTGCCAGCACATCGTTAAGTGGATCATTAAGTACAAGCATTTCAG 3505

Qy        654 spCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerValPheAspThrA 674
```

```
Db       3506 ATTCAACGAGTACGTCGACATCCGACAGTGCATCAACGTCAACGAGTGAGAGTGAGTCAG 3565

Qy        674 snSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuA 694

Db       3566 ACAGTACAAGTACGTCATTA---AGTGAATCGACAAGCACAAGTCTTTCGGACTCAACAA 3622

Qy        694 laLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnT 714

Db       3623 GTACGTCAACGTCAGAAAGTGCATCAACGTCAACGAGTGAGAGTGACTCAACAAGTGAAA 3682

Qy        714 hrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisH 734

Db       3683 GTACGTCATTAAGCGAGTCGACAAGCACAAGCGTTTCAGATTCAACAAGCGCGTCGACAT 3742

Qy        734 isAsp----------AspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgL 750

Db       3743 CCGACAGTGCATCAACATCAACGAGTGTGAGCGACTCGGAAAGTGCAAGTACATCCATAA 3802

Qy        750 ysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaA 770

Db       3803 GTGAATCATTAAGTACGAGCGTTTCAGATTCAACGAGTACGTCGACATCCGACAGTGCAT 3862

Qy        770 laThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisV 790

Db       3863 CAACGTCAACGAGTGAGAGTGACTCAACAAGTGAAAGTACGTCATTAAGCGAGTCGATAA 3922

Qy        790 alMetGlySerSerSerAsn-TyrAlaSerArgGlnProValIleAlaProLeuAspArg 809

Db       3923 GCACAAGCGTTTCAGATTCAACAAGCGCGTCGACATCCGACAGTGCGTCCACATCAACGA 3982

Qy        810 TyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThr 829

Db       3983 GTGAGAGTGAGTCAGACAGTGCAAGTACAT---CGTTA-AGTGGCTCGACAAGTACGAGT 4038

Qy        830 MetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProLys 849

Db       4039 CTTTCAGATTCAACGAGTACATCGACATCAGACAGTGCATCAACGTCAACAAGTGAGAGC 4098

Qy        850 GluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhe 869

Db       4099 GACTCGGAAAGAGCAAGTACATCCTTAAGTGGCTCGACAAGTACGAGTCTTTCAGATTCA 4158

Qy        870 ProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyr 889

Db       4159 ACGAGTACATCGACATCAGACAGTGCATCAACATCAACGAGTGTGAGTGACTCAAACAGT 4218

Qy        890 AlaHisAsnGlnTyrLysGlySerThrSerThrSerTyr---------- 902

Db       4219 GCCAGCACATCGTTAAGTGGCTCATTAAGTACAAGCGTTTCAGACTCAACAAGTACATCA 4278

Qy        903 ----------GlySerAsn 905

Db       4279 ACATCCGATAGTGCATCTGCATCAACAAGTGAGAGCGACTCGGAAAGAGCAAGTACATCG 4338

Qy        906 LeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArg-HisGlnLeuHisAspLe 925

Db       4339 TTGAGTGGCTCAACAAGTACAAGCATTTCAGACTCAACAAGTACATCAA---------- 4387

Qy        925 uHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGl 945

Db       4388 -CATCAGACAGTGCATCAACATCAACGAGTGTGAGTGAG-TCAAACAGTACAAGCACATC 4445

Qy        945 uIleAla----------AsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLy 960

Db       4446 CATAAGTGAATCATTAAGTACGAGCGTTTCAGATTCAACGAGTACATCGACATCAGACAG 4505

Qy        960 sLeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAl 980

Db       4506 TGCATCAACATCAACGAGTGTGAGTGACTCGGACAGTGCAAGTACGTCATCAAGTGAATC 4565

Qy        980 aLeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSe 1000

Db       4566 AGTAAGCACAAGT------GATTCAGAATCAACAAGCACATCCACATCAGATAGTGCATC 4619

Qy       1000 rSerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGl 1020

Db       4620 CACATCAACGAGTGTGAGTGAGTCAAACAGTACAAGCACATCATTGAGTGGCTCAACAAG 4679

Qy       1020 yLysMetValHisProLeuAspArgPhe------ValArgGlnAspIleCysIle-ThrA 1038

Db       4680 TACAAGCGTTTCAGA-CTCAACAAGTACATCGACATCAGACAGTGCATCTGCATCAACAA 4738

Qy       1038 snLysAsnProAlaAspPheThrThrIleSerAsn 1049

Db       4739 GTGAGAGCGACTCGGACAGTGCAAGTACGTCATCA 4773

RESULT 8
US-09-134-001C-1685/c
; Sequence 1685, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1685
;   LENGTH: 3561
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1685

Alignment Scores:
Pred. No.:                1.9e-08        Length:          3561
Score:                    197.50         Matches:         185
Percent Similarity:       34.6%          Conservative:    154
Best Local Similarity:    18.9%          Mismatches:      450
Query Match:              3.6%           Indels:          193
DB:                       3              Gaps:            31

US-09-828-068-2 (1-1057) x US-09-134-001C-1685 (1-3561)

Qy        117 ThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAsp 136

Db       2738 ACATCCGACAGTGCATCA---------ACGTCAACAAGCGAGAGTGACTCAACAAGTGAA 2688

Qy        137 GlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGly--------- 153

Db       2687 AGTACATCGTTGAGTGAATCATTAAGCACAAGCGTTTCAGATTCAACAAGCGCGTCAACG 2628

Qy        154 SerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSer--- 172

Db       2627 TCAGAAAGTGCATCCACATCAACGAGTGAGAGTGAGTCAAACAGTGCCAGCACATCGTTA 2568

Qy        173 ------ThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSer 190

Db       2567 AGTGGATCATTAAGTACAAGCATTTCAGAT---TCAACGAGTACGTCGACATCCGACAGT 2511

Qy        191 GlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyPro 210

Db       2510 GCATCAACGTCAACGAGTGAGAGTGAGTCAGACAGTACAAGTACGTCATTAAGT------ 2457

Qy        211 AlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGly 230

Db       2456 ------------GAATCGACAAGCACAAGTCTTTCGGACTCAACAAGTACGTCAACG--- 2412

Qy        231 AlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnPro 250

Db       2412 ------------------------------------------------------------ 2412
```


Qy 251 ProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLys 270
Db 2411 ----------TCAGAAAGTGCATCAACGTCAACGAGTGAGAGTGACTCAACA 2370
Qy 271 ThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSer 290
Db 2369 AGTGAAAGTACGTCATTAAGCGAGTCGACAAGC---ACAAGCGTTTCAGATTCAACAAGC 2313
Qy 291 GlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProVal 310
Db 2312 GCGTCGACATCCGACAGTGCATCAACATCAACGAGTGTGAGC---GACTCGGAAAGTGCA 2256
Qy 311 SerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArg 330
Db 2255 AGTACATCCATAAGTGAATCATTAAGTACGAGCGTTTCAGATTCAACGAGTACGTCGACA 2196
Qy 331 SerSerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLys 350
Db 2195 TCCGACAGTGCATCAACGTCAACGAGTGAGAGTGACTCAACAAGTGAAAGTACGTCATTA 2136
Qy 351 LysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLys 370
Db 2135 AGCGAGTCGATAAGCACAAGCGTTTCAGATTCA---------- 2103
Qy 371 ValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGlu 390
Db 2102 ----------ACAAGCGCGTCGACATCCGACAGTGCGTCC------ 2073
Qy 391 ValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValPro 410
Db 2072 ----------ACATCAACGAGTGAGAGTGAGTCAGACAGTGCAAGTACATCGTTAAGT 2025
Qy 411 MetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLys 430
Db 2024 GGCTCGACAAGTACGAGTCTTTCAGATTCAACGAGTACATCGACATCAGACAGTGCATCA 1965
Qy 431 SerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeu 450
Db 1964 ACGTCAACAAGTGAGAGCGACTCGGAAAGAGCAAGTACATCCTTAAGTGGCTCGACA--- 1908
Qy 451 MetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHis 470
Db 1908 ---------- 1908
Qy 471 ProAlaGlyAsnLeuSerAsnLysLysValThrProThrAla----------Ser 485
Db 1907 ---AGTACGAGTCTTTCAGATTCAACGAGTACATCGACATCAGACAGTGCATCAACATCA 1851
Qy 486 ThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr 505
Db 1850 ACGAGTGTGAGTGACTCAAACAGTGCCAGCACATCGTTAAGTGGCTCATTAAGTACAAGC 1791
Qy 506 AspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThr 525
Db 1790 ----------GTTTCAGAC---TCAACAAGTACATCAACATCCGATAGTGCATCT 1749
Qy 526 AlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSer 545
Db 1748 GCATCAACAAGTGAGAGCGACTCGGAAAGAGCAAGTACATCGTTGAGTGGCTCAACAAGT 1689
Qy 546 ThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThr 565
Db 1688 ACA----------AGCATTTCAGACTCAACAAGTACATCAACA 1656
Qy 566 GluAsnSerValLeuSerHisSerAlaLysValSerProAlaGluHisAspIleGlnIle 585
Db 1655 TCAGACAGTGCA---TCAACATCAACGAGTGTGAGTGAGTCAAACAGTACAAGCACATCC 1599
Qy 586 MetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysLeuGluVal 605
Db 1598 ATAAGTGAATCATTAAGTACGAGC----------GTT 1572
Qy 606 ThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAla 625
Db 1571 TCAGATTCAACGAGTACATCGACATCAGAC---------- 1542
Qy 626 LysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIle 645
Db 1541 ---AGTGCATCAACATCAACGAGTGTGAGTGACTCGGACAGTGCAAGTACGTCATCAAGT 1485
Qy 646 GlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAsp 665
Db 1484 GAATCAGTAAGCACAAGTGATTCAGAATCAACAAGCACATCCACATCAGATAGTGCATCC 1425
Qy 666 TyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThr 685
Db 1424 ACATCAACGAGTGTGAGTGAGTCAAACAGTACAAGCACATCATTG---AGTGGCTCAACA 1368
Qy 686 GlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsn 705
Db 1367 AGTACAAGCGTTTCAGACTCAACAAGTACATCGACATCAGACAGTGCATCTGCATCAACA 1308
Qy 706 PheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAla 725
Db 1307 AGTGAGAGCGACTCGGACAGTGCAAGTACGTCATCAAGTGAATCAGTAAGCACAAGCGTT 1248
Qy 726 AlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGlu 745
Db 1247 TCAGATTCAACAAGCGCGTCA----------ACGTCAGAAAGTGCGTCAACA 1206
Qy 746 HisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db 1205 TCAACGAGTGTGAGTGACTCGAAT----------AGTGCAAGTACGTCA 1167
Qy 766 AsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeu 785
Db 1166 TTAAGTGAATCGACAAGTACGAGTCTTTCAGATTCAACGAGTATGTCGACATCAGACAGT 1107
Qy 786 ThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAla 805
Db 1106 GCGTCAACATCAACAAGTGAGAGCGACTCGGACAGTGCAAGTACATCG------TTAAGT 1053
Qy 806 ProLeuAspArgTyrAlaGluArg----------AlaValAsnGln-Va 818
Db 1052 GA-CTCGACAAGTAC---GAGCGTTTCAGAATCGACAAGTACATCAACATCAACGAGTGT 997
Qy 818 lHisAlaArgAsnPheProSerThr---IleAlaThrMetGluAlaSerLysLeuCysAs 837
Db 996 GAGTGCCTCAAACAGTACAAGTACGTCATTAAGTGACTCGAGAAGTACGAGTCTTTCAGA 937
Qy 837 pArgArgAsn----------AlaGlyGlnValValLeuTyrProLysGluSerMe 852
Db 936 CTCAACGAGCACATCGACATCAGAAAGTGGATCAACGTCAACAAGTGAGAGCGACTCGGA 877
Qy 852 tProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheProAsnTy 872
Db 876 CAGTGCAAGTACATCCTTAAGTGAGTCAACAAGTACAAGCATTTCAGACTCAACGAGTAC 817
Qy 872 rGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrAlaHisAs 892
Db 816 ATCAACATCAGATAGTGCATCTACATCTATGAGTGTAAGCGACTCAAACAGAGCAAGTAC 757
Qy 892 nGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysIleProLe 912
Db 756 GTCCTTAAGTGACTCGACAAGCACAAGCGTTTCAGATTCAACAAGCGCGTCA-------- 705
Qy 912 uThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeuArgProHi 932
Db 704 -ACATCAGAAAGCGCATCA----------ACGTCAACAAGAGAGAG 670
Qy 932 sProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAs 952
Db 669 TGAATCAACAAGTGCAAGTACGTCATTGAGTGAATCGACAAGCACAAGCGTTTCAGATTC 610
Qy 952 nCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleThrSerHisGlnMe 972

```
Db        609 AACAAGTACATCGACATCCGACAGTGCGTCAACGTCAACAAGTGAGAGTGACTCAAACAG 550

Qy        972 tAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaLysTrpAsnAl 992

Db        549 TGAAAGTACATCGTTAAGTGAGTCGACAAGCACAAGCGTT---TCAGATTCAACAAGCGC 493

Qy        992 aLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaArgAsnSerIleAl 1012

Db        492 GTCAACGTCAGCAAGTGCATCAACATCAACGAGC---GTGAGTGACTCCAATAGCGCAAG 436

Qy       1012 aGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspArgPheValArgGl 1032

Db        435 TACGTCATTAAGTGGCTCAACAAGTACAAGCGTTTCAGA-CTCAACGAGTACGTCGACAT 377

Qy       1032 nAspIleCysIle-ThrAsnLysAsnProAlaAsp---------------------- 1043

Db        376 CAGCAAGTGCATCAACATCAACGAGCGAGAGCGACTCCGATAGCGCAAGTACGTCGTTAA 317

Qy       1044 -----------PheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057

Db        316 GTGGCTCAACAAGTACAAGCATTTCAGACTC-AACGAGTACGTCGACATCAGA 265
```

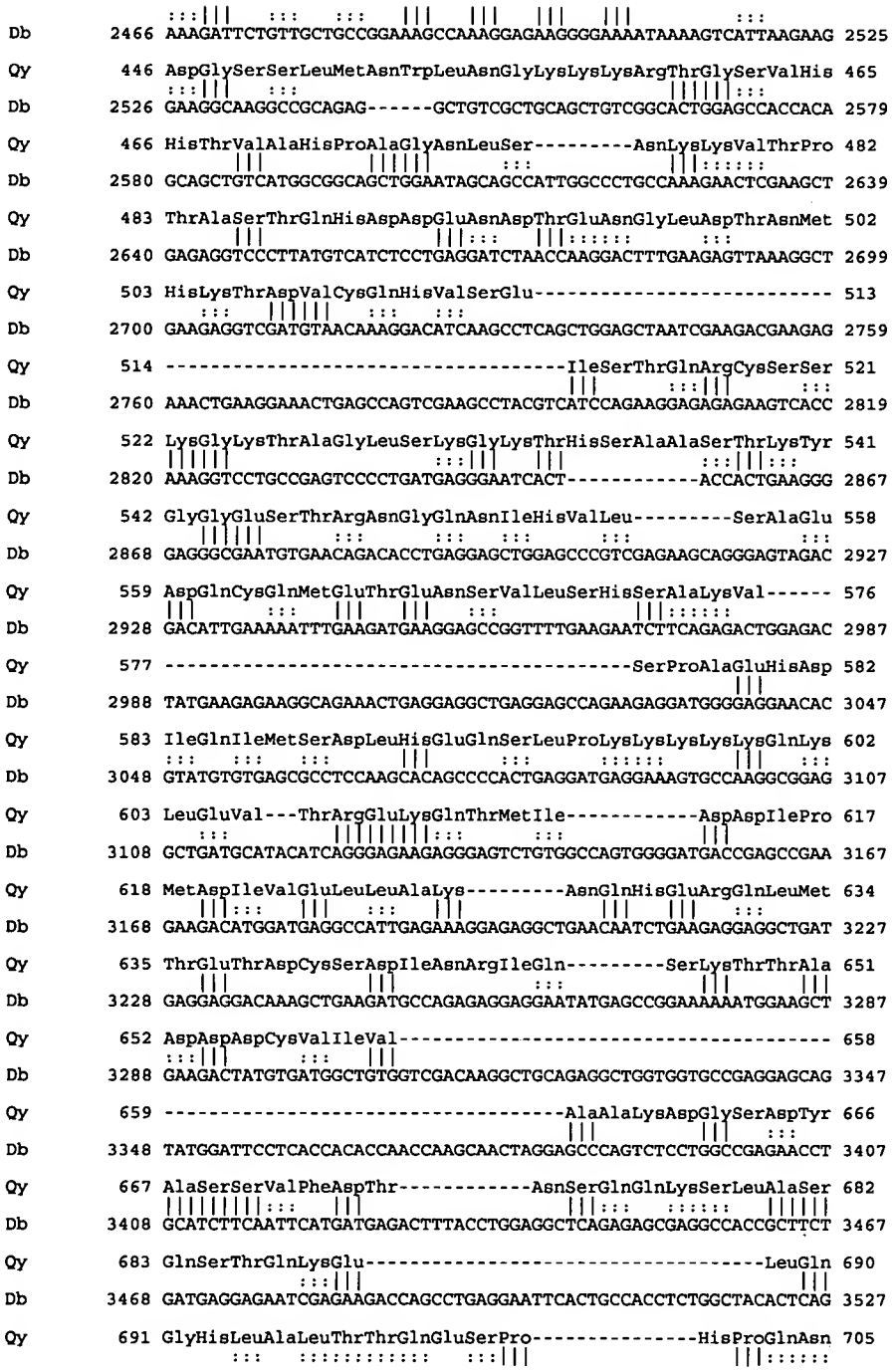
RESULT 9
US-09-949-016-4366
; Sequence 4366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4366
; LENGTH: 7571
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4366

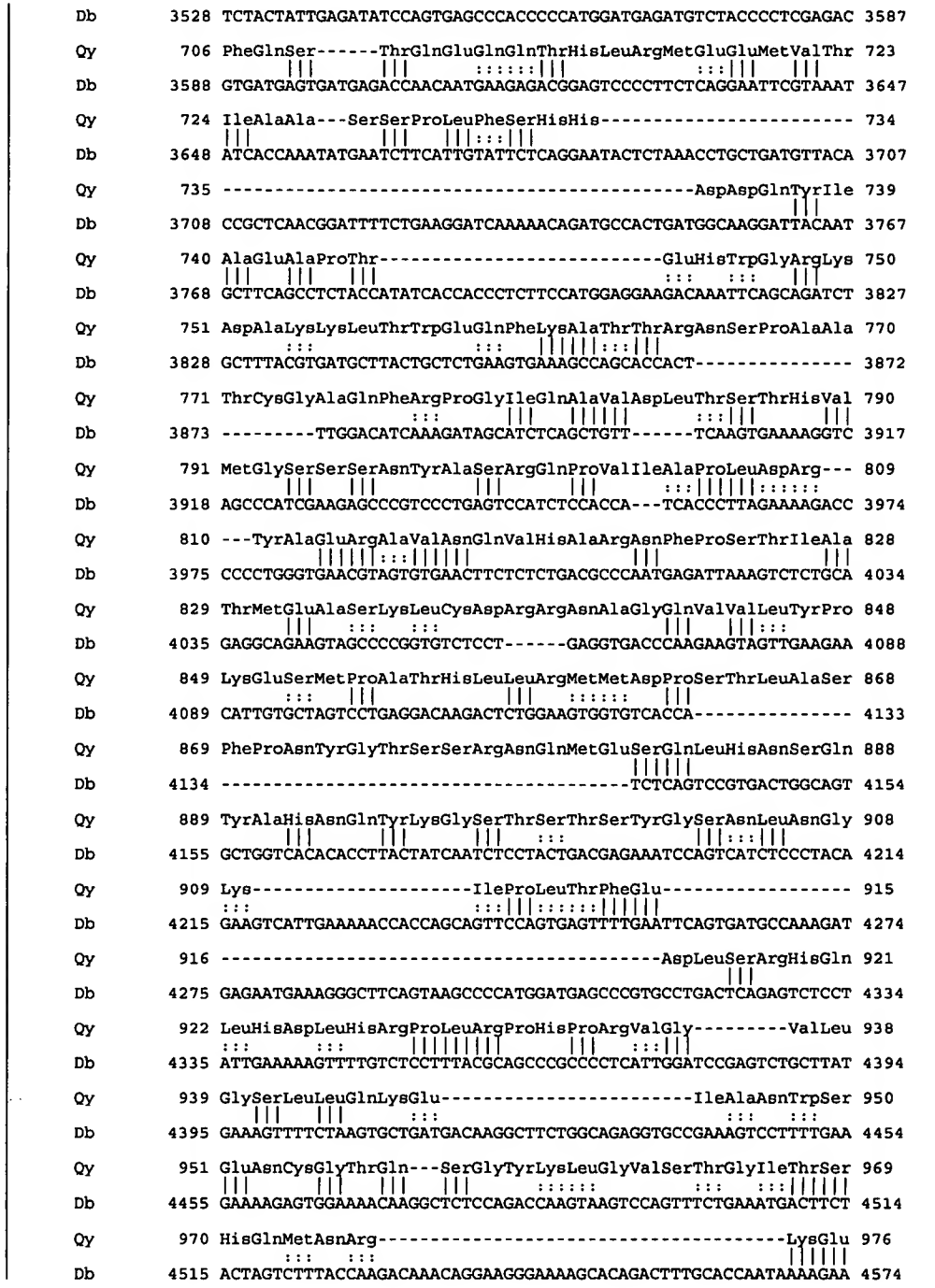
Alignment Scores:

Pred. No.:	8.98e-08	Length:	7571
Score:	196.50	Matches:	239
Percent Similarity:	33.8%	Conservative:	189
Best Local Similarity:	18.9%	Mismatches:	490
Query Match:	3.6%	Indels:	351
DB:	3	Gaps:	54

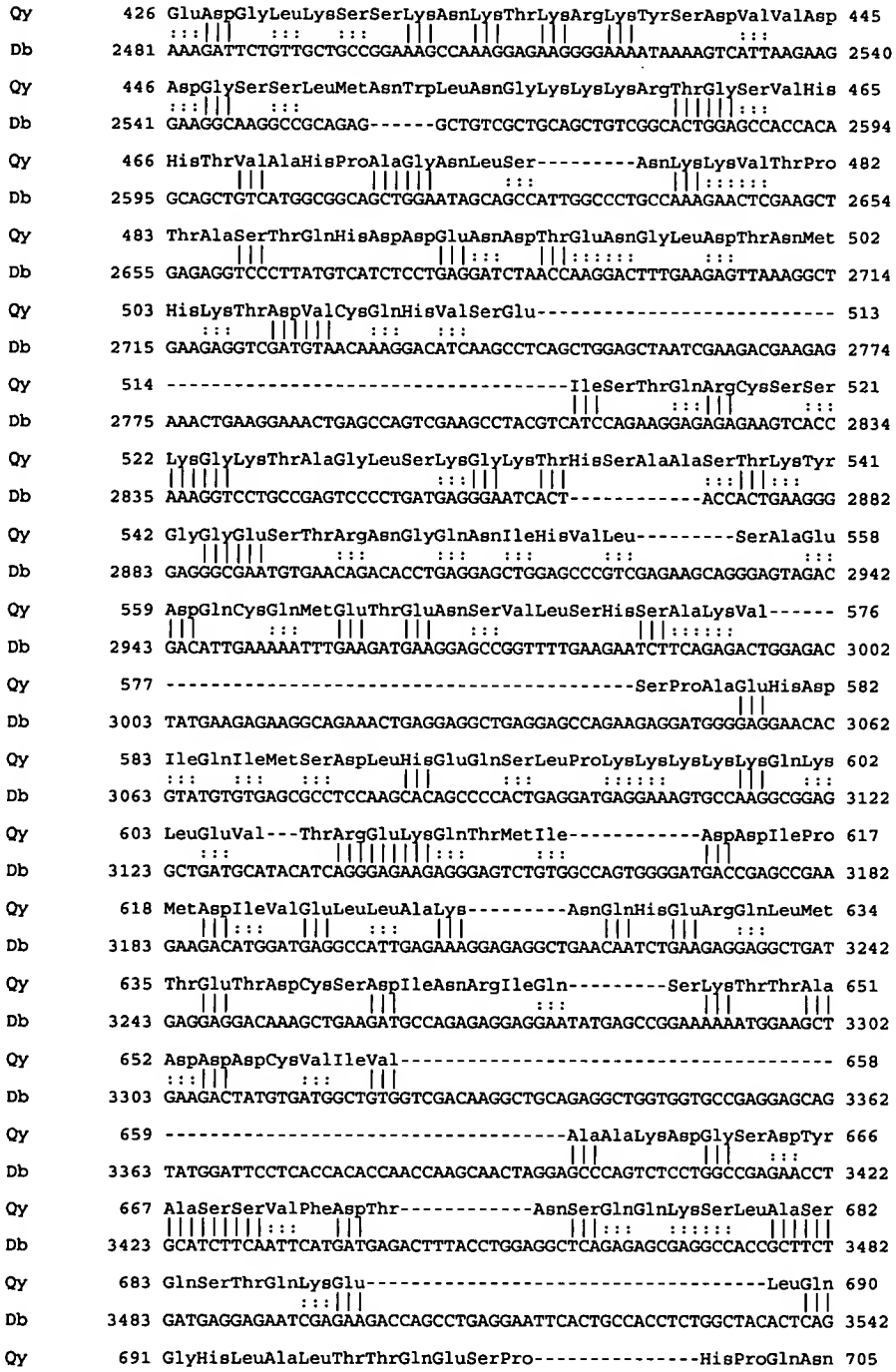
US-09-828-068-2 (1-1057) x US-09-949-016-4366 (1-7571)

```
Qy         52 AlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeuGlnLysLys 71

Db       1271 GCCAAACATCAAGATGAAGAGAAGCATAGAAGA---------AGCCTGCTTCACTCTCCA 1321

Qy         72 AspProLysPheCysSerLeuSerArgIlePheHisAspGlnLys-----------Lys 87

Db       1322 GTACCTAAA--------------CAAATTGTCCATGAAACCAGAACCTCTGTTTAGAAG 1366

Qy         88 CysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAsp 107

Db       1367 TGTAGGCAATACTATTGATCCTG-TCATTCTTTTCC-----------AAAAAATGGGAG 1413

Qy        107 ------------------------------------------------------------ 107

Db       1414 TAGGTAAACTTGAGATGTATGTGCTTAATCCAGTCAAGAGCAGCAAGGAAATGCAGTATT 1473

Qy         108 ---CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThr 126

Db       1474 TTATGCAGCAGTGGACTGGTACCAACAAAGACAAGGCTGAATTCATT-CTGCCT------ 1526

Qy         127 LeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSer--- 145

Db       1527 --------------AATGGTCAAGAAGTAGATCTCCCGATTTCCTACTTAACTTCAGTC 1571

Qy         146 -----------ThrPheValProAlaSerValGlySerGlnLysVal----------- 157

Db       1572 TCATCTTTGATTGTGTGGCATCCAGCAAACCCTGCGGAGAAAATCATCCGAGTCCTGTTT 1631

Qy         158 ---SerProSerThrGlnSerSer------GlnGly-----------LysAsnAlaAsp 170

Db       1632 CCTGGGAACAGCACCCAGTACAACATCCTGGAAGGGTTGGAAAAGCTCAAACATCTAGAC 1691

Qy         171 ArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSer 190

Db       1692 TTTCTGAAGCAGCCACTGGCCACCCAAAAGGATCTCACTGGCCAGGTGCCCACTCCTGTG 1751

Qy         191 GlyLysAsnGlyAlaAlaGluAlaAsnThrAspSer------ProMetLysAspLeuGln 208

Db       1752 GTGAAACAAACAAAACTGAAACAGAGGGCTGATAGCCGAGAAAGTCTGAAGCCAGCCGCA 1811

Qy         209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228

Db       1812 AAACCACTTCCTAGCAAATCCGTGCGCAAGGAGTCAAAAGAA---------------- 1853

Qy         229 ValGlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAsp 248

Db       1854 --------------GAAACCCCTGAGGTCACA---------AAAGTGAATCACGTGGAA 1889

Qy         249 GlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsn 268

Db       1890 AAGCCACCCAAAGTTGAAAGCAAAGAAAAGGTAATGGTGAAAAAAGACAAGCCAGTAAAA 1949

Qy         269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288

Db       1950 ACAGAGACCAAACCTTCAGTGACTGAAAAGGAGGTTCCCAGCAAAGAAGAGCCATCTCCA 2009

Qy         289 MetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys 308

Db       2010 GTG-----------AAAGCCGAGGTGGCTGAGAAG------CAAGCCACAGATGTCAAA 2051

Qy         309 ProValSerGly--------------GlnLysCysGluGlnIleCysAsnGluProCys 323

Db       2052 CCCAAAGCTGCCAAGGAGAAGACGGTGAAAAAGGAAACAAAGGTAAAGCCTGAAGACAAG 2111

Qy         324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343

Db       2112 AAAGAGGAGAAAGAAAAGCCAAAGAAAGAAGTGGCTAAAAAGGAGGACAAAACACCTATC 2171

Qy         344 LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspVal-------------- 358

Db       2172 AAGAAGGAGGAAAAACCAAAAAAGGAAGAGGTGAAAAAAGAAGTCAAAAAAGAGATCAAG 2231

Qy         359 ---SerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIle 377

Db       2232 AAAGAAGAGAAAAAAGAACCCAAGAAAGAGGTTAAGAAAGAAACACCGCCAAAGGAAGTC 2291

Qy         378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397

Db       2292 ------AAGAAGGAAGTTAAGAAGGAAGAGAAGAAGGAAGTGAAAAAGGAAGAAAAGGAA 2345

Qy         398 Asp--------------------ProCysGluAspAspArgSerThrIleProVal--- 409

Db       2346 CCCAAAAAAGAAATTAAGAAGCTCCCTAAAGACGCAAAGAAATCATCTACTCCTCTGTCT 2405

Qy         410 ------------ProMetGluValSerMetAspIleProValSerAsnHisThrValGly 425

Db       2406 GAAGCAAAAAAACCAGCTGCTTTAAAACCAAAAGTACCCAAGAAGGAAGAGTCTGTCAAG 2465

Qy         426 GluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAsp 445
```


Db 2466 AAAGATTCTGTTGCTGCCGGAAAGCCAAAGGAGAAGGGGAAAATAAAAGTCATTAAGAAG 2525
Qy 446 AspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHis 465
Db 2526 GAAGGCAAGGCCGCAGAG------GCTGTCGCTGCAGCTGTCGGCACTGGAGCCACCACA 2579
Qy 466 HisThrValAlaHisProAlaGlyAsnLeuSer---------AsnLysLysValThrPro 482
Db 2580 GCAGCTGTCATGGCGGCAGCTGGAATAGCAGCCATTGGCCCTGCCAAAGAACTCGAAGCT 2639
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMet 502
Db 2640 GAGAGGTCCCTTATGTCATCTCCTGAGGATCTAACCAAGGACTTTGAAGAGTTAAAGGCT 2699
Qy 503 HisLysThrAspValCysGlnHisValSerGlu-------------------------- 513
Db 2700 GAAGAGGTCGATGTAACAAAGGACATCAAGCCTCAGCTGGAGCTAATCGAAGACGAAGAG 2759
Qy 514 ----------------------------------IleSerThrGlnArgCysSerSer 521
Db 2760 AAACTGAAGGAAACTGAGCCAGTCGAAGCCTACGTCATCCAGAAGGAGAGAGAAGTCACC 2819
Qy 522 LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyr 541
Db 2820 AAAGGTCCTGCCGAGTCCCCTGATGAGGGAATCACT-----------ACCACTGAAGGG 2867
Qy 542 GlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeu---------SerAlaGlu 558
Db 2868 GAGGGCGAATGTGAACAGACACCTGAGGAGCTGGAGCCCGTCGAGAAGCAGGGAGTAGAC 2927
Qy 559 AspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysVal------ 576
Db 2928 GACATTGAAAAATTTGAAGATGAAGGAGCCGGTTTTGAAGAATCTTCAGAGACTGGAGAC 2987
Qy 577 --SerProAlaGluHisAsp 582
Db 2988 TATGAAGAGAAGGCAGAAACTGAGGAGGCTGAGGAGCCAGAAGAGGATGGGGAGGAACAC 3047
Qy 583 IleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLys 602
Db 3048 GTATGTGTGAGCGCCTCCAAGCACAGCCCCACTGAGGATGAGGAAAGTGCCAAGGCGGAG 3107
Qy 603 LeuGluVal---ThrArgGluLysGlnThrMetIle-----------AspAspIlePro 617
Db 3108 GCTGATGCATACATCAGGGAGAAGAGGGAGTCTGTGGCCAGTGGGGATGACCGAGCCGAA 3167
Qy 618 MetAspIleValGluLeuLeuAlaLys---------AsnGlnHisGluArgGlnLeuMet 634
Db 3168 GAAGACATGGATGAGGCCATTGAGAAAGGAGAGGCTGAACAATCTGAAGAGGAGGCTGAT 3227
Qy 635 ThrGluThrAspCysSerAspIleAsnArgIleGln---------SerLysThrThrAla 651
Db 3228 GAGGAGGACAAAGCTGAAGATGCCAGAGAGGAGGAATATGAGCCGGAAAAAATGGAAGCT 3287
Qy 652 AspAspAspCysValIleVal-- 658
Db 3288 GAAGACTATGTGATGGCTGTGGTCGACAAGGCTGCAGAGGCTGGTGGTGCCGAGGAGCAG 3347
Qy 659 ----------------------------------AlaAlaLysAspGlySerAspTyr 666
Db 3348 TATGGATTCCTCACCACACCAACCAAGCAACTAGGAGCCCAGTCTCCTGGCCGAGAACCT 3407
Qy 667 AlaSerSerValPheAspThr------------AsnSerGlnGlnLysSerLeuAlaSer 682
Db 3408 GCATCTTCAATTCATGATGAGACTTTACCTGGAGGCTCAGAGAGCGAGGCCACCGCTTCT 3467
Qy 683 GlnSerThrGlnLysGlu----------------------------------LeuGln 690
Db 3468 GATGAGGAGAATCGAGAAGACCAGCCTGAGGAATTCACTGCCACCTCTGGCTACACTCAG 3527
Qy 691 GlyHisLeuAlaLeuThrThrGlnGluSerPro--------------HisProGlnAsn 705


Db 3528 TCTACTATTGAGATATCCAGTGAGCCCACCCCCATGGATGAGATGTCTACCCCTCGAGAC 3587
Qy 706 PheGlnSer------ThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThr 723
Db 3588 GTGATGAGTGATGAGACCAACAATGAAGAGACGGAGTCCCCTTCTCAGGAATTCGTAAAT 3647
Qy 724 IleAlaAla---SerSerProLeuPheSerHisHis---------------------- 734
Db 3648 ATCACCAAATATGAATCTTCATTGTATTCTCAGGAATACTCTAAACCTGCTGATGTTACA 3707
Qy 735 --AspAspGlnTyrIle 739
Db 3708 CCGCTCAACGGATTTTCTGAAGGATCAAAAACAGATGCCACTGATGGCAAGGATTACAAT 3767
Qy 740 AlaGluAlaProThr--------------------------GluHisTrpGlyArgLys 750
Db 3768 GCTTCAGCCTCTACCATATCACCACCCTCTTCCATGGAGGAAGACAAATTCAGCAGATCT 3827
Qy 751 AspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAla 770
Db 3828 GCTTTACGTGATGCTTACTGCTCTGAAGTGAAAGCCAGCACCACT--------------- 3872
Qy 771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
Db 3873 ---------TTGGACATCAAAGATAGCATCTCAGCTGTT------TCAAGTGAAAAGGTC 3917
Qy 791 MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArg--- 809
Db 3918 AGCCCATCGAAGAGCCCGTCCCTGAGTCCATCTCCACCA---TCACCCTTAGAAAAGACC 3974
Qy 810 ---TyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAla 828
Db 3975 CCCCTGGGTGAACGTAGTGTGAACTTCTCTCTGACGCCCAATGAGATTAAAGTCTCTGCA 4034
Qy 829 ThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPro 848
Db 4035 GAGGCAGAAGTAGCCCCGGTGTCTCCT------GAGGTGACCCAAGAAGTAGTTGAAGAA 4088
Qy 849 LysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSer 868
Db 4089 CATTGTGCTAGTCCTGAGGACAAGACTCTGGAAGTGGTGTCACCA--------------- 4133
Qy 869 PheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGln 888
Db 4134 ------------------------------------TCTCAGTCCGTGACTGGCAGT 4154
Qy 889 TyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGly 908
Db 4155 GCTGGTCACACACCTTACTATCAATCTCCTACTGACGAGAAATCCAGTCATCTCCCTACA 4214
Qy 909 Lys---------------------IleProLeuThrPheGlu----------------- 915
Db 4215 GAAGTCATTGAAAAACCACCAGCAGTTCCAGTGAGTTTTGAATTCAGTGATGCCAAAGAT 4274
Qy 916 --AspLeuSerArgHisGln 921
Db 4275 GAGAATGAAAGGGCTTCAGTAAGCCCCATGGATGAGCCCGTGCCTGACTCAGAGTCTCCT 4334
Qy 922 LeuHisAspLeuHisArgProLeuArgProHisProArgValGly---------ValLeu 938
Db 4335 ATTGAAAAAGTTTTGTCTCCTTTACGCAGCCCGCCCCTCATTGGATCCGAGTCTGCTTAT 4394
Qy 939 GlySerLeuLeuGlnLysGlu----------------------IleAlaAsnTrpSer 950
Db 4395 GAAAGTTTTCTAAGTGCTGATGACAAGGCTTCTGGCAGAGGTGCCGAAAGTCCTTTTGAA 4454
Qy 951 GluAsnCysGlyThrGln---SerGlyTyrLysLeuGlyValSerThrGlyIleThrSer 969
Db 4455 GAAAAGAGTGGAAAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTTCTGAAATGACTTCT 4514
Qy 970 HisGlnMetAsnArg-------------------------------------LysGlu 976
Db 4515 ACTAGTCTTTACCAAGACAAACAGGAAGGGAAAAGCACAGACTTTGCACCAATAAAAGAA 4574

```
Qy       977 HisPhe--------------------------GluAlaLeuAsnSerGlyMetPheSer 987
Db      4575 GACTTTGGCCAAGAAAAGAAAACTGATGATGTTGAAGCCATGAGTTCTCAACCAGCACTG 4634

Qy       988 AlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSer------------AlaAsp 1003
Db      4635 GCT---CTGGATGAAAGGAAATTAGGAGATGTTTCTCCCACACAAATAGATGTCAGTCAG 4691

Qy      1004 PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGly------LysGlyLys 1021
Db      4692 TTTGGATCTTTTAAAGAAGACACTAAGATGTCCATTTCTGAAGGTACTGTCTCAGACAAG 4751

Qy      1022 MetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
Db      4752 TCAGCTACTCCTGTTGATGAGGGCGTAGCAGAAGACACGTACTCTCATATGGAGGGTGTG 4811

Qy      1042 AlaAspPheThrThrIleSer 1048
Db      4812 GCCTCAGTGTCCACAGCCTCA 4832

RESULT 10
US-09-976-594-725
; Sequence 725, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 725
;   LENGTH: 8146
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 4295277CB1
;   NAME/KEY: unsure
;   LOCATION: 8083
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-725

Alignment Scores:
Pred. No.:              1.02e-07      Length:         8146
Score:                  196.50        Matches:        239
Percent Similarity:     33.8%         Conservative:   189
Best Local Similarity:  18.9%         Mismatches:     490
Query Match:            3.6%          Indels:         351
DB:                     3             Gaps:           54

US-09-828-068-2 (1-1057) x US-09-976-594-725 (1-8146)

Qy         52 AlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysLys 71
Db       1286 GCCAAACATCAAGATGAAGAGAAGCATAGAAGA---------AGCCTGCTTCACTCTCCA 1336

Qy         72 AspProLysPheCysSerLeuSerArgIlePheHisAspGlnLys------------Lys 87
Db       1337 GTACCTAAA--------------CAAATTGTCCATGAAACCAGAACCTCTGTTTAGAAG 1381

Qy         88 CysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAsp 107
Db       1382 TGTAGGCAATACTATTGATCCTG-TCATTCTTTTCC------------AAAAAATGGGAG 1428

Qy        107 ------------------------------------------------------------ 107
Db       1429 TAGGTAAACTTGAGATGTATGTGCTTAATCCAGTCAAGAGCAGCAAGGAAATGCAGTATT 1488

Qy        108 ---CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThr 126
Db       1489 TTATGCAGCAGTGGACTGGTACCAACAAAGACAAGGCTGAATTCATT-CTGCCT------ 1541

Qy        127 LeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSer--- 145
Db       1542 --------------AATGGTCAAGAAGTAGATCTCCCGATTTCCTACTTAACTTCAGTC 1586

Qy        146 ------------ThrPheValProAlaSerValGlySerGlnLysVal------------ 157
Db       1587 TCATCTTTGATTGTGTGGCATCCAGCAAACCCTGCGGAGAAAATCATCCGAGTCCTGTTT 1646

Qy        158 ---SerProSerThrGlnSerSer------GlnGly------------LysAsnAlaAsp 170
Db       1647 CCTGGGAACAGCACCCAGTACAACATCCTGGAAGGGTTGGAAAAGCTCAAACATCTAGAC 1706

Qy        171 ArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSer 190
Db       1707 TTTCTGAAGCAGCCACTGGCCACCCAAAAGGATCTCACTGGCCAGGTGCCCACTCCTGTG 1766

Qy        191 GlyLysAsnGlyAlaAlaGluAlaAsnThrAspSer------ProMetLysAspLeuGln 208
Db       1767 GTGAAACAAACAAAACTGAAACAGAGGGCTGATAGCCGAGAAAGTCTGAAGCCAGCCGCA 1826

Qy        209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
Db       1827 AAACCACTTCCTAGCAAATCCGTGCGCAAGGAGTCAAAAGAA---------------- 1868

Qy        229 ValGlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAsp 248
Db       1869 --------------GAAACCCCTGAGGTCACA---------AAAGTGAATCACGTGGAA 1904

Qy        249 GlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsn 268
Db       1905 AAGCCACCCAAAGTTGAAAGCAAAGAAAAGGTAATGGTGAAAAAAGACAAGCCAGTAAAA 1964

Qy        269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
Db       1965 ACAGAGACCAAACCTTCAGTGACTGAAAAGGAGGTTCCCAGCAAAGAAGAGCCATCTCCA 2024

Qy        289 MetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys 308
Db       2025 GTG-----------AAAGCCGAGGTGGCTGAGAAG------CAAGCCACAGATGTCAAA 2066

Qy        309 ProValSerGly--------------GlnLysCysGluGlnIleCysAsnGluProCys 323
Db       2067 CCCAAAGCTGCCAAGGAGAAGACGGTGAAAAAGGAAACAAAGGTAAAGCCTGAAGACAAG 2126

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
Db       2127 AAAGAGGAGAAAGAAAAGCCAAAGAAAGAAGTGGCTAAAAAGGAGGACAAAACACCTATC 2186

Qy        344 LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspVal--------------- 358
Db       2187 AAGAAGGAGGAAAAACCAAAAAAGGAAGAGGTGAAAAAAGAAGTCAAAAAAGAGATCAAG 2246

Qy        359 ---SerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIle 377
Db       2247 AAAGAAGAGAAAAAAGAACCCAAGAAAGAGGTTAAGAAAGAAACACCGCCAAAGGAAGTC 2306

Qy        378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397
Db       2307 ------AAGAAGGAAGTTAAGAAGGAAGAGAAGAAGGAAGTGAAAAAGGAAGAAAAGGAA 2360

Qy        398 Asp--------------------ProCysGluAspAspArgSerThrIleProVal--- 409
Db       2361 CCCAAAAAAGAAATTAAGAAGCTCCCTAAAGACGCAAAGAAATCATCTACTCCTCTGTCT 2420

Qy        410 ------------ProMetGluValSerMetAspIleProValSerAsnHisThrValGly 425
Db       2421 GAAGCAAAAAAACCAGCTGCTTTAAAACCAAAAGTACCCAAGAAGGAAGAGTCTGTCAAG 2480
```


Qy 426 GluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAsp 445
Db 2481 AAAGATTCTGTTGCTGCCGGAAAGCCAAAGGAGAAGGGGAAAATAAAAGTCATTAAGAAG 2540
Qy 446 AspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHis 465
Db 2541 GAAGGCAAGGCCGCAGAG------GCTGTCGCTGCAGCTGTCGGCACTGGAGCCACCACA 2594
Qy 466 HisThrValAlaHisProAlaGlyAsnLeuSer---------AsnLysLysValThrPro 482
Db 2595 GCAGCTGTCATGGCGGCAGCTGGAATAGCAGCCATTGGCCCTGCCAAAGAACTCGAAGCT 2654
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMet 502
Db 2655 GAGAGGTCCCTTATGTCATCTCCTGAGGATCTAACCAAGGACTTTGAAGAGTTAAAGGCT 2714
Qy 503 HisLysThrAspValCysGlnHisValSerGlu-------------------------- 513
Db 2715 GAAGAGGTCGATGTAACAAAGGACATCAAGCCTCAGCTGGAGCTAATCGAAGACGAAGAG 2774
Qy 514 ----------------------------------IleSerThrGlnArgCysSerSer 521
Db 2775 AAACTGAAGGAAACTGAGCCAGTCGAAGCCTACGTCATCCAGAAGGAGAGAGAAGTCACC 2834
Qy 522 LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyr 541
Db 2835 AAAGGTCCTGCCGAGTCCCCTGATGAGGGAATCACT-----------ACCACTGAAGGG 2882
Qy 542 GlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeu---------SerAlaGlu 558
Db 2883 GAGGGCGAATGTGAACAGACACCTGAGGAGCTGGAGCCCGTCGAGAAGCAGGGAGTAGAC 2942
Qy 559 AspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysVal------ 576
Db 2943 GACATTGAAAAATTTGAAGATGAAGGAGCCGGTTTTGAAGAATCTTCAGAGACTGGAGAC 3002
Qy 577 --SerProAlaGluHisAsp 582
Db 3003 TATGAAGAGAAGGCAGAAACTGAGGAGGCTGAGGAGCCAGAAGAGGATGGGGAGGAACAC 3062
Qy 583 IleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLys 602
Db 3063 GTATGTGTGAGCGCCTCCAAGCACAGCCCCACTGAGGATGAGGAAAGTGCCAAGGCGGAG 3122
Qy 603 LeuGluVal---ThrArgGluLysGlnThrMetIle-----------AspAspIlePro 617
Db 3123 GCTGATGCATACATCAGGGAGAAGAGGGAGTCTGTGGCCAGTGGGGATGACCGAGCCGAA 3182
Qy 618 MetAspIleValGluLeuLeuAlaLys---------AsnGlnHisGluArgGlnLeuMet 634
Db 3183 GAAGACATGGATGAGGCCATTGAGAAAGGAGAGGCTGAACAATCTGAAGAGGAGGCTGAT 3242
Qy 635 ThrGluThrAspCysSerAspIleAsnArgIleGln---------SerLysThrThrAla 651
Db 3243 GAGGAGGACAAAGCTGAAGATGCCAGAGAGGAGGAATATGAGCCGGAAAAAATGGAAGCT 3302
Qy 652 AspAspAspCysValIleVal-- 658
Db 3303 GAAGACTATGTGATGGCTGTGGTCGACAAGGCTGCAGAGGCTGGTGGTGCCGAGGAGCAG 3362
Qy 659 ----------------------------------AlaAlaLysAspGlySerAspTyr 666
Db 3363 TATGGATTCCTCACCACACCAACCAAGCAACTAGGAGCCCAGTCTCCTGGCCGAGAACCT 3422
Qy 667 AlaSerSerValPheAspThr-----------AsnSerGlnGlnLysSerLeuAlaSer 682
Db 3423 GCATCTTCAATTCATGATGAGACTTTACCTGGAGGCTCAGAGAGCGAGGCCACCGCTTCT 3482
Qy 683 GlnSerThrGlnLysGlu---------------------------------LeuGln 690
Db 3483 GATGAGGAGAATCGAGAAGACCAGCCTGAGGAATTCACTGCCACCTCTGGCTACACTCAG 3542
Qy 691 GlyHisLeuAlaLeuThrThrGlnGluSerPro--------------HisProGlnAsn 705
Db 3543 TCTACTATTGAGATATCCAGTGAGCCCACCCCCATGGATGAGATGTCTACCCCTCGAGAC 3602
Qy 706 PheGlnSer------ThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThr 723
Db 3603 GTGATGAGTGATGAGACCAACAATGAAGAGACGGAGTCCCCTTCTCAGGAATTCGTAAAT 3662
Qy 724 IleAlaAla---SerSerProLeuPheSerHisHis------------------------ 734
Db 3663 ATCACCAAATATGAATCTTCATTGTATTCTCAGGAATACTCTAAACCTGCTGATGTTACA 3722
Qy 735 --AspAspGlnTyrIle 739
Db 3723 CCGCTCAACGGATTTTCTGAAGGATCAAAAACAGATGCCACTGATGGCAAGGATTACAAT 3782
Qy 740 AlaGluAlaProThr--------------------------GluHisTrpGlyArgLys 750
Db 3783 GCTTCAGCCTCTACCATATCACCACCCTCTTCCATGGAGGAAGACAAATTCAGCAGATCT 3842
Qy 751 AspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAla 770
Db 3843 GCTTTACGTGATGCTTACTGCTCTGAAGTGAAAGCCAGCACCACT--------------- 3887
Qy 771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
Db 3888 ---------TTGGACATCAAAGATAGCATCTCAGCTGTT------TCAAGTGAAAAGGTC 3932
Qy 791 MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArg--- 809
Db 3933 AGCCCATCGAAGAGCCCGTCCCTGAGTCCATCTCCACCA---TCACCCTTAGAAAAGACC 3989
Qy 810 ---TyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnProSerThrIleAla 828
Db 3990 CCCCTGGGTGAACGTAGTGTGAACTTCTCTCTGACGCCCAATGAGATTAAAGTCTCTGCA 4049
Qy 829 ThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPro 848
Db 4050 GAGGCAGAAGTAGCCCCGGTGTCTCCT------GAGGTGACCCAAGAAGTAGTTGAAGAA 4103
Qy 849 LysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSer 868
Db 4104 CATTGTGCTAGTCCTGAGGACAAGACTCTGGAAGTGGTGTCACCA--------------- 4148
Qy 869 PheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGln 888
Db 4149 ------------------------------------TCTCAGTCCGTGACTGGCAGT 4169
Qy 889 TyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGly 908
Db 4170 GCTGGTCACACACCTTACTATCAATCTCCTACTGACGAGAAATCCAGTCATCTCCCTACA 4229
Qy 909 Lys--------------------IleProLeuThrPheGlu------------------ 915
Db 4230 GAAGTCATTGAAAAACCACCAGCAGTTCCAGTGAGTTTTGAATTCAGTGATGCCAAAGAT 4289
Qy 916 --AspLeuSerArgHisGln 921
Db 4290 GAGAATGAAAGGGCTTCAGTAAGCCCCATGGATGAGCCCGTGCCTGACTCAGAGTCTCCT 4349
Qy 922 LeuHisAspLeuHisArgProLeuArgProHisProArgValGly---------ValLeu 938
Db 4350 ATTGAAAAAGTTTTGTCTCCTTTACGCAGCCCGCCCCTCATTGGATCCGAGTCTGCTTAT 4409
Qy 939 GlySerLeuLeuGlnLysGlu----------------------IleAlaAsnTrpSer 950
Db 4410 GAAAGTTTTCTAAGTGCTGATGACAAGGCTTCTGGCAGAGGTGCCGAAAGTCCTTTTGAA 4469
Qy 951 GluAsnCysGlyThrGln---SerGlyTyrLysLeuGlyValSerThrGlyIleThrSer 969
Db 4470 GAAAAGAGTGGAAAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTTCTGAAATGACTTCT 4529
Qy 970 HisGlnMetAsnArg-------------------------------------LysGlu 976

```
Db       4530 ACTAGTCTTTACCAAGACAAACAGGAAGGGAAAAGCACAGACTTTGCACCAATAAAAGAA 4589

Qy        977 HisPhe-------------------------GluAlaLeuAsnSerGlyMetPheSer 987
                 |||                         ||||||::::::|||
Db       4590 GACTTTGGCCAAGAAAAGAAAACTGATGATGTTGAAGCCATGAGTTCTCAACCAGCACTG 4649

Qy        988 AlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSer------------AlaAsp 1003
              |||      :::      :::||||||   ||||||   :::           :::
Db       4650 GCT---CTGGATGAAAGGAAATTAGGAGATGTTTCTCCCACACAAATAGATGTCAGTCAG 4706

Qy       1004 PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGly------LysGlyLys 1021
              |||   |||   :::               |||   :::   |||             |||
Db       4707 TTTGGATCTTTTAAAGAAGACACTAAGATGTCCATTTCTGAAGGTACTGTCTCAGACAAG 4766

Qy       1022 MetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
                    |||:::|||       |||   :::|||                 :::
Db       4767 TCAGCTACTCCTGTTGATGAGGGCGTAGCAGAAGACACGTACTCTCATATGGAGGGTGTG 4826

Qy       1042 AlaAspPheThrThrIleSer 1048
              |||      :::|||   |||
Db       4827 GCCTCAGTGTCCACAGCCTCA 4847
```

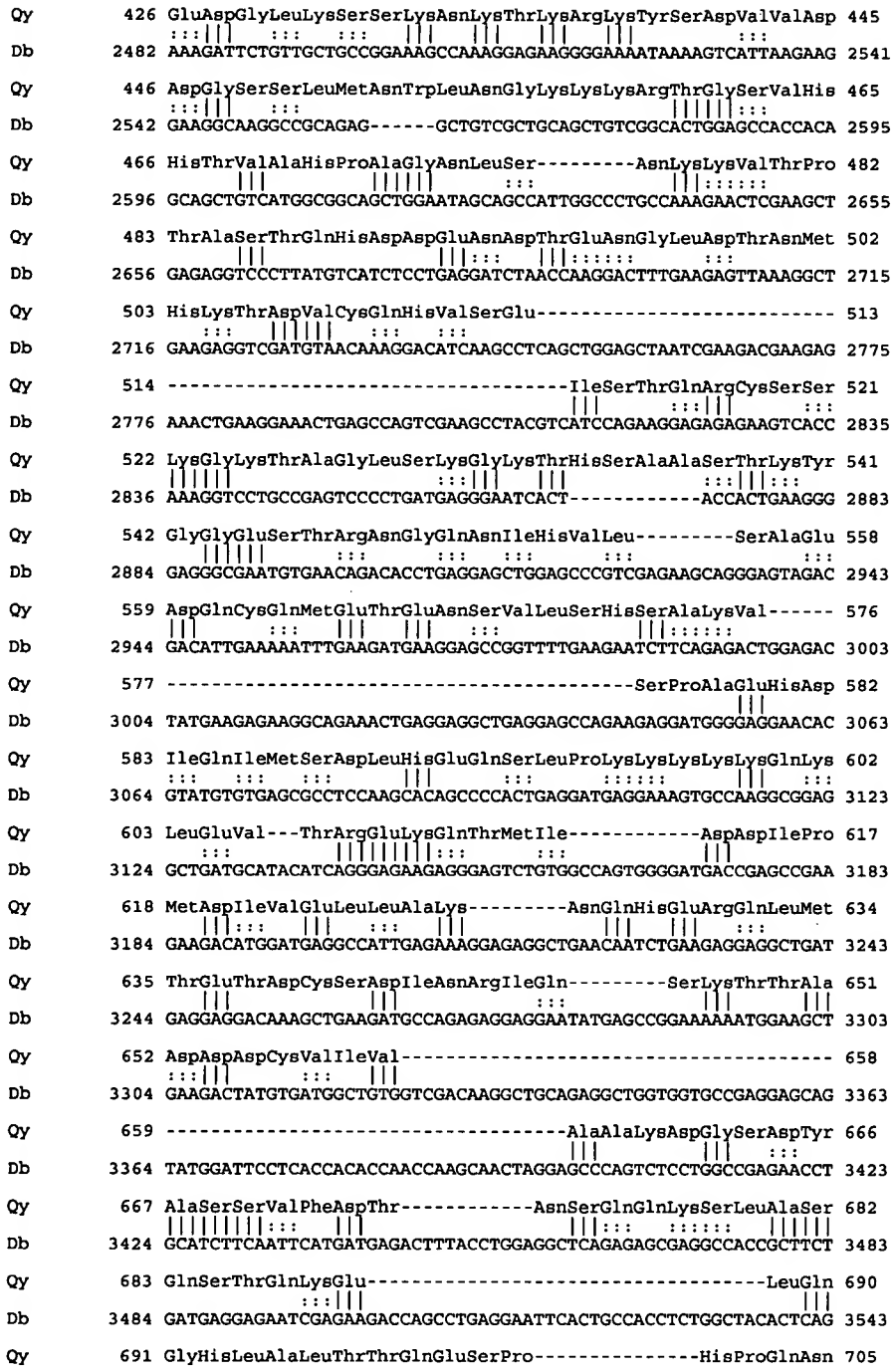
```
RESULT 11
US-09-566-921-32
; Sequence 32, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
;  APPLICANT:  Loring, Jeanne F.
;  APPLICANT:  Tingley, Debora W.
;  APPLICANT:  Edwards, Carla M.
;  TITLE OF INVENTION:  GENES EXPRESSED IN ALZHEIMER'S DISEASE
;  FILE REFERENCE:  PA-0024 US
;  CURRENT APPLICATION NUMBER: US/09/566,921
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  138
;  SOFTWARE:  PERL Program
; SEQ ID NO 32
;   LENGTH: 11917
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6682888 475473.1
;   NAME/KEY: unsure
;   LOCATION: 9384, 10290-10321, 10325
;   OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-32
```

Alignment Scores:			
Pred. No.:	1.99e-07	Length:	11917
Score:	196.50	Matches:	239
Percent Similarity:	33.8%	Conservative:	189
Best Local Similarity:	18.9%	Mismatches:	490
Query Match:	3.6%	Indels:	351
DB:	3	Gaps:	54

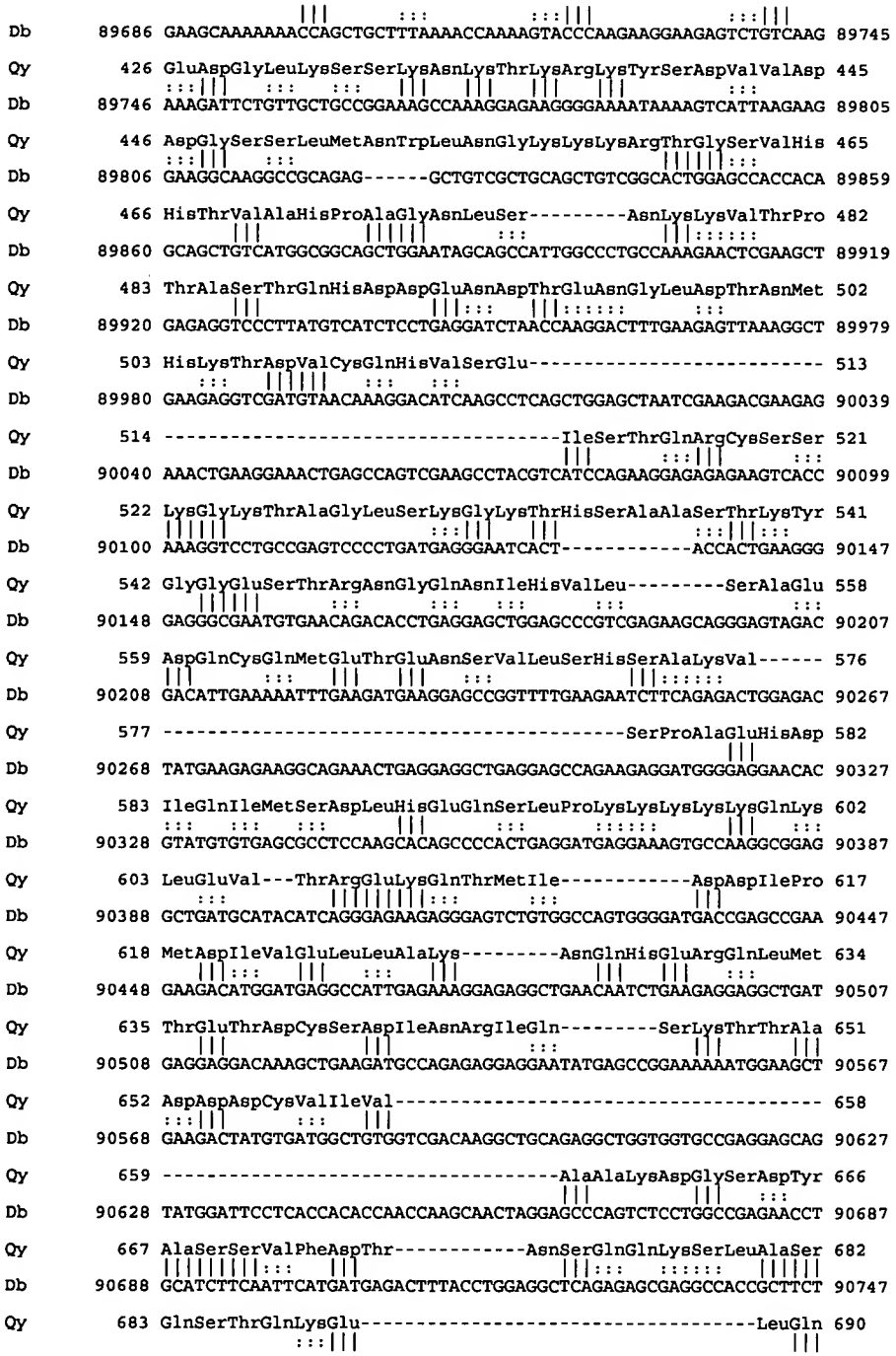
US-09-828-068-2 (1-1057) x US-09-566-921-32 (1-11917)

```
Qy         52 AlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysLys 71
              |||:::||||||   |||             |||         :::|||||||
Db       1287 GCCAAACATCAAGATGAAGAGAAGCATAGAAGA---------AGCCTGCTTCACTCTCCA 1337

Qy         72 AspProLysPheCysSerLeuSerArgIlePheHisAspGlnLys-----------Lys 87
                 ||||||                ::|||   |||:::   :::            |||
Db       1338 GTACCTAAA--------------CAAATTGTCCATGAAACCAGAACCTCTGTTTAGAAG 1382

Qy         88 CysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAsp 107
              |||   ::::::      |||   |||   ||||||           :::::::|||:::
Db       1383 TGTAGGCAATACTATTGATCCTG-TCATTCTTTTCC-----------AAAAAATGGGAG 1429

Qy        107 ------------------------------------------------------------ 107

Db       1430 TAGGTAAACTTGAGATGTATGTGCTTAATCCAGTCAAGAGCAGCAAGGAAATGCAGTATT 1489
```

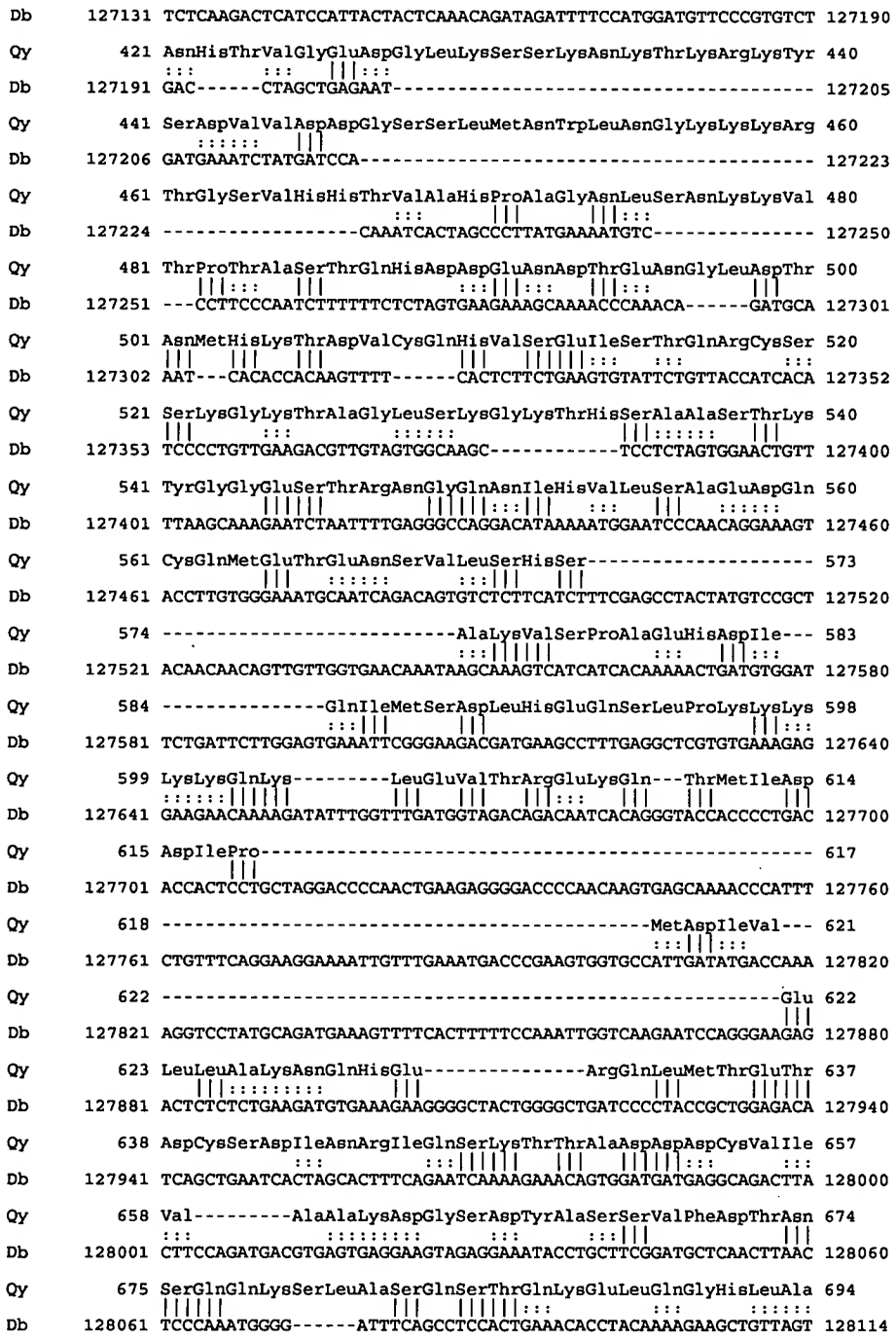
```
Qy        108 ---CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThr 126
                  ||||||      |||         ||||||      |||         |||
Db       1490 TTATGCAGCAGTGGACTGGTACCAACAAAGACAAGGCTGAATTCATT-CTGCCT------ 1542

Qy        127 LeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSer--- 145
                       ||||||                ||:::::::::   |||
Db       1543 --------------AATGGTCAAGAAGTAGATCTCCCGATTTCCTACTTAACTTCAGTC 1587

Qy        146 ------------ThrPheValProAlaSerValGlySerGlnLysVal------------ 157
                           :::   ||||||:::         :::   :::
Db       1588 TCATCTTTGATTGTGTGGCATCCAGCAAACCCTGCGGAGAAAATCATCCGAGTCCTGTTT 1647

Qy        158 ---SerProSerThrGlnSerSer------GlnGly------------LysAsnAlaAsp 170
                    |||||||||   :::      :::|||            |||:::   |||
Db       1648 CCTGGGAACAGCACCCAGTACAACATCCTGGAAGGGTTGGAAAAGCTCAAACATCTAGAC 1707

Qy        171 ArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSer 190
                       |||   :::   |||:::            :::         |||
Db       1708 TTTCTGAAGCAGCCACTGGCCACCCAAAAGGATCTCACTGGCCAGGTGCCCACTCCTGTG 1767

Qy        191 GlyLysAsnGlyAlaAlaGluAlaAsnThrAspSer------ProMetLysAspLeuGln 208
                 |||           :::         ||||||         :::|||
Db       1768 GTGAAACAAACAAAACTGAAACAGAGGGCTGATAGCCGAGAAAGTCTGAAGCCAGCCGCA 1827

Qy        209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
                 |||      :::      |||             |||
Db       1828 AAACCACTTCCTAGCAAATCCGTGCGCAAGGAGTCAAAAGAA----------------- 1869

Qy        229 ValGlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAsp 248
                       |||   |||::::::|||          :::|||||||      :::
Db       1870 --------------GAAACCCCTGAGGTCACA---------AAAGTGAATCACGTGGAA 1905

Qy        249 GlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsn 268
              :::||||||                 :::|||::::::|||:::::::::
Db       1906 AAGCCACCCAAAGTTGAAAGCAAAGAAAAGGTAATGGTGAAAAAAGACAAGCCAGTAAAA 1965

Qy        269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
                 :::|||:::   ::::::      ::::::           ::::::|||:::|||
Db       1966 ACAGAGACCAAACCTTCAGTGACTGAAAAGGAGGTTCCCAGCAAAGAAGAGCCATCTCCA 2025

Qy        289 MetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys 308
                 :::            :::   :::||||||||||:::                |||   |||
Db       2026 GTG-----------AAAGCCGAGGTGGCTGAGAAG------CAAGCCACAGATGTCAAA 2067

Qy        309 ProValSerGly--------------GlnLysCysGluGlnIleCysAsnGluProCys 323
                 |||   :::                 :::|||      ::::::      |||
Db       2068 CCCAAAGCTGCCAAGGAGAAGACGGTGAAAAAGGAAACAAAGGTAAAGCCTGAAGACAAG 2127

Qy        324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
                 :::|||      |||         :::      :::|||   ||||||      :::
Db       2128 AAAGAGGAGAAAGAAAAGCCAAAGAAAGAAGTGGCTAAAAAGGAGGACAAAACACCTATC 2187

Qy        344 LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspVal--------------- 358
                 ||||||::::::      ||||||          :::   :::|||
Db       2188 AAGAAGGAGGAAAAACCAAAAAAGGAAGAGGTGAAAAAAGAAGTCAAAAAAGAGATCAAG 2247

Qy        359 ---SerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIle 377
                    :::   |||      :::::::::   ||||||           |||:::
Db       2248 AAAGAAGAGAAAAAAGAACCCAAGAAAGAGGTTAAGAAAGAAACACCGCCAAAGGAAGTC 2307

Qy        378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397
                    :::|||:::                ||||||   :::|||
Db       2308 ------AAGAAGGAAGTTAAGAAGGAAGAGAAGAAGGAAGTGAAAAAGGAAGAAAAGGAA 2361

Qy        398 Asp--------------------ProCysGluAspAspArgSerThrIleProVal--- 409
                                     |||   :::      :::|||:::   |||:::
Db       2362 CCCAAAAAAGAAATTAAGAAGCTCCCTAAAGACGCAAAGAAATCATCTACTCCTCTGTCT 2421

Qy        410 ------------ProMetGluValSerMetAspIleProValSerAsnHisThrValGly 425
                           |||      :::      :::|||            :::|||
Db       2422 GAAGCAAAAAAACCAGCTGCTTTAAAACCAAAAGTACCCAAGAAGGAAGAGTCTGTCAAG 2481
```

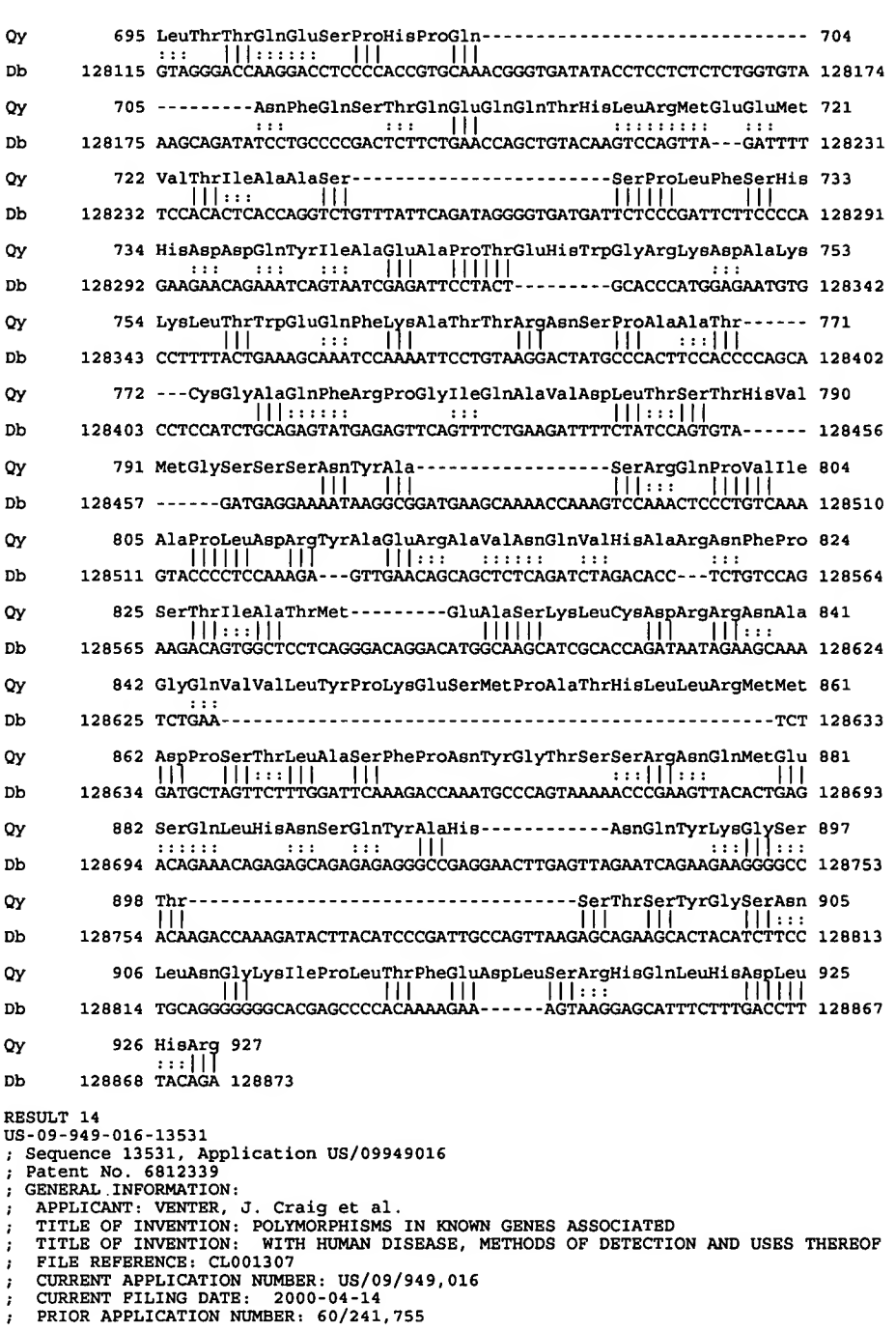

Qy 426 GluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAsp 445
Db 2482 AAAGATTCTGTTGCTGCCGGAAAGCCAAAGGAGAAGGGGAAAATAAAAGTCATTAAGAAG 2541
Qy 446 AspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHis 465
Db 2542 GAAGGCAAGGCCGCAGAG------GCTGTCGCTGCAGCTGTCGGCACTGGAGCCACCACA 2595
Qy 466 HisThrValAlaHisProAlaGlyAsnLeuSer---------AsnLysLysValThrPro 482
Db 2596 GCAGCTGTCATGGCGGCAGCTGGAATAGCAGCCATTGGCCCTGCCAAAGAACTCGAAGCT 2655
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMet 502
Db 2656 GAGAGGTCCCTTATGTCATCTCCTGAGGATCTAACCAAGGACTTTGAAGAGTTAAAGGCT 2715
Qy 503 HisLysThrAspValCysGlnHisValSerGlu-------------------------- 513
Db 2716 GAAGAGGTCGATGTAACAAAGGACATCAAGCCTCAGCTGGAGCTAATCGAAGACGAAGAG 2775
Qy 514 ----------------------------------IleSerThrGlnArgCysSerSer 521
Db 2776 AAACTGAAGGAAACTGAGCCAGTCGAAGCCTACGTCATCCAGAAGGAGAGAGAAGTCACC 2835
Qy 522 LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyr 541
Db 2836 AAAGGTCCTGCCGAGTCCCCTGATGAGGGAATCACT-----------ACCACTGAAGGG 2883
Qy 542 GlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeu---------SerAlaGlu 558
Db 2884 GAGGGCGAATGTGAACAGACACCTGAGGAGCTGGAGCCCGTCGAGAAGCAGGGAGTAGAC 2943
Qy 559 AspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysVal------ 576
Db 2944 GACATTGAAAAATTTGAAGATGAAGGAGCCGGTTTTGAAGAATCTTCAGAGACTGGAGAC 3003
Qy 577 --SerProAlaGluHisAsp 582
Db 3004 TATGAAGAGAAGGCAGAAACTGAGGAGGCTGAGGAGCCAGAAGAGGATGGGGAGGAACAC 3063
Qy 583 IleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLys 602
Db 3064 GTATGTGTGAGCGCCTCCAAGCACAGCCCCACTGAGGATGAGGAAAGTGCCAAGGCGGAG 3123
Qy 603 LeuGluVal---ThrArgGluLysGlnThrMetIle-----------AspAspIlePro 617
Db 3124 GCTGATGCATACATCAGGGAGAAGAGGGAGTCTGTGGCCAGTGGGGATGACCGAGCCGAA 3183
Qy 618 MetAspIleValGluLeuLeuAlaLys---------AsnGlnHisGluArgGlnLeuMet 634
Db 3184 GAAGACATGGATGAGGCCATTGAGAAAGGAGAGGCTGAACAATCTGAAGAGGAGGCTGAT 3243
Qy 635 ThrGluThrAspCysSerAspIleAsnArgIleGln---------SerLysThrThrAla 651
Db 3244 GAGGAGGACAAAGCTGAAGATGCCAGAGAGGAGGAATATGAGCCGGAAAAAATGGAAGCT 3303
Qy 652 AspAspAspCysValIleVal-- 658
Db 3304 GAAGACTATGTGATGGCTGTGGTCGACAAGGCTGCAGAGGCTGGTGGTGCCGAGGAGCAG 3363
Qy 659 ----------------------------------AlaAlaLysAspGlySerAspTyr 666
Db 3364 TATGGATTCCTCACCACACCAACCAAGCAACTAGGAGCCCAGTCTCCTGGCCGAGAACCT 3423
Qy 667 AlaSerSerValPheAspThr-----------AsnSerGlnGlnLysSerLeuAlaSer 682
Db 3424 GCATCTTCAATTCATGATGAGACTTTACCTGGAGGCTCAGAGAGCGAGGCCACCGCTTCT 3483
Qy 683 GlnSerThrGlnLysGlu---------------------------------LeuGln 690
Db 3484 GATGAGGAGAATCGAGAAGACCAGCCTGAGGAATTCACTGCCACCTCTGGCTACACTCAG 3543
Qy 691 GlyHisLeuAlaLeuThrThrGlnGluSerPro--------------HisProGlnAsn 705
Db 3544 TCTACTATTGAGATATCCAGTGAGCCCACCCCCATGGATGAGATGTCTACCCCTCGAGAC 3603
Qy 706 PheGlnSer------ThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThr 723
Db 3604 GTGATGAGTGATGAGACCAACAATGAAGAGACGGAGTCCCCTTCTCAGGAATTCGTAAAT 3663
Qy 724 IleAlaAla---SerSerProLeuPheSerHisHis----------------------- 734
Db 3664 ATCACCAAATATGAATCTTCATTGTATTCTCAGGAATACTCTAAACCTGCTGATGTTACA 3723
Qy 735 --AspAspGlnTyrIle 739
Db 3724 CCGCTCAACGGATTTTCTGAAGGATCAAAAACAGATGCCACTGATGGCAAGGATTACAAT 3783
Qy 740 AlaGluAlaProThr--------------------------GluHisTrpGlyArgLys 750
Db 3784 GCTTCAGCCTCTACCATATCACCACCCTCTTCCATGGAGGAAGACAAATTCAGCAGATCT 3843
Qy 751 AspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAla 770
Db 3844 GCTTTACGTGATGCTTACTGCTCTGAAGTGAAAGCCAGCACCACT--------------- 3888
Qy 771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
Db 3889 ---------TTGGACATCAAAGATAGCATCTCAGCTGTT------TCAAGTGAAAAGGTC 3933
Qy 791 MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArg--- 809
Db 3934 AGCCCATCGAAGAGCCCGTCCCTGAGTCCATCTCCACCA---TCACCCTTAGAAAAGACC 3990
Qy 810 ---TyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAla 828
Db 3991 CCCCTGGGTGAACGTAGTGTGAACTTCTCTCTGACGCCCAATGAGATTAAAGTCTCTGCA 4050
Qy 829 ThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPro 848
Db 4051 GAGGCAGAAGTAGCCCCGGTGTCTCCT------GAGGTGACCCAAGAAGTAGTTGAAGAA 4104
Qy 849 LysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSer 868
Db 4105 CATTGTGCTAGTCCTGAGGACAAGACTCTGGAAGTGGTGTCACCA--------------- 4149
Qy 869 PheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGln 888
Db 4150 ---------------------------------------TCTCAGTCCGTGACTGGCAGT 4170
Qy 889 TyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGly 908
Db 4171 GCTGGTCACACACCTTACTATCAATCTCCTACTGACGAGAAATCCAGTCATCTCCCTACA 4230
Qy 909 Lys---------------------IleProLeuThrPheGlu------------------ 915
Db 4231 GAAGTCATTGAAAAACCACCAGCAGTTCCAGTGAGTTTTGAATTCAGTGATGCCAAAGAT 4290
Qy 916 --AspLeuSerArgHisGln 921
Db 4291 GAGAATGAAAGGGCTTCAGTAAGCCCCATGGATGAGCCCGTGCCTGACTCAGAGTCTCCT 4350
Qy 922 LeuHisAspLeuHisArgProLeuArgProHisProArgValGly---------ValLeu 938
Db 4351 ATTGAAAAAGTTTTGTCTCCTTTACGCAGCCCGCCCCTCATTGGATCCGAGTCTGCTTAT 4410
Qy 939 GlySerLeuLeuGlnLysGlu----------------------IleAlaAsnTrpSer 950
Db 4411 GAAAGTTTTCTAAGTGCTGATGACAAGGCTTCTGGCAGAGGTGCCGAAAGTCCTTTTGAA 4470
Qy 951 GluAsnCysGlyThrGln---SerGlyTyrLysLeuGlyValSerThrGlyIleThrSer 969
Db 4471 GAAAAGAGTGGAAAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTTCTGAAATGACTTCT 4530
Qy 970 HisGlnMetAsnArg-----------------------------------LysGlu 976

```
Db       4531 ACTAGTCTTTACCAAGACAAACAGGAAGGGAAAAGCACAGACTTTGCACCAATAAAAGAA 4590

Qy        977 HisPhe--------------------------GluAlaLeuAsnSerGlyMetPheSer 987
                 |||                          ||||||:::::::|||
Db       4591 GACTTTGGCCAAGAAAAGAAAACTGATGATGTTGAAGCCATGAGTTCTCAACCAGCACTG 4650

Qy        988 AlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSer------------AlaAsp 1003
                |||      :::      :::||||||   ||||||   :::            :::
Db       4651 GCT---CTGGATGAAAGGAAATTAGGAGATGTTTCTCCCACACAAATAGATGTCAGTCAG 4707

Qy       1004 PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGly------LysGlyLys 1021
                |||   |||   :::              |||   :::   |||           |||
Db       4708 TTTGGATCTTTTAAAGAAGACACTAAGATGTCCATTTCTGAAGGTACTGTCTCAGACAAG 4767

Qy       1022 MetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
                     |||:::|||       |||   :::|||                 :::
Db       4768 TCAGCTACTCCTGTTGATGAGGGCGTAGCAGAAGACACGTACTCTCATATGGAGGGTGTG 4827

Qy       1042 AlaAspPheThrThrIleSer 1048
                |||      :::|||   |||
Db       4828 GCCTCAGTGTCCACAGCCTCA 4848

RESULT 12
US-09-949-016-16108
; Sequence 16108, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16108
;   LENGTH: 101300
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(101300)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16108

Alignment Scores:
Pred. No.:              8.52e-06        Length:         101300
Score:                  196.50          Matches:        239
Percent Similarity:     33.8%           Conservative:   189
Best Local Similarity:  18.9%           Mismatches:     490
Query Match:            3.6%            Indels:         351
DB:                     3               Gaps:           54

US-09-828-068-2 (1-1057) x US-09-949-016-16108 (1-101300)

Qy          52 AlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysLys 71
               |||:::||||||   |||           |||         :::|||||||
Db       88551 GCCAAACATCAAGATGAAGAGAAGCATAGAAGA---------AGCCTGCTTCACTCTCCA 88601

Qy          72 AspProLysPheCysSerLeuSerArgIlePheHisAspGlnLys------------Lys 87
                  ||||||               :::|||   |||:::   :::            |||
Db       88602 GTACCTAAA--------------CAAATTGTCCATGAAACCAGAACCTCTGTTTAGAAG 88646

Qy          88 CysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAsp 107
               |||   ::::::     |||   |||   ||||||           :::::::|||:::
Db       88647 TGTAGGCAATACTATTGATCCTG-TCATTCTTTTCC------------AAAAAATGGGAG 88693

Qy         107 ------------------------------------------------------------ 107

Db       88694 TAGGTAAACTTGAGATGTATGTGCTTAATCCAGTCAAGAGCAGCAAGGAAATGCAGTATT 88753

Qy         108 ---CysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThr 126
                  ||||||     |||         ||||||      |||         |||
Db       88754 TTATGCAGCAGTGGACTGGTACCAACAAAGACAAGGCTGAATTCATT-CTGCCT------ 88806

Qy         127 LeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSer--- 145
                        ||||||                ||||||||||    |||
Db       88807 --------------AATGGTCAAGAAGTAGATCTCCCGATTTCCTACTTAACTTCAGTC 88851

Qy         146 -----------ThrPheValProAlaSerValGlySerGlnLysVal------------ 157
                         :::   ||||||:::         :::   :::
Db       88852 TCATCTTTGATTGTGTGGCATCCAGCAAACCCTGCGGAGAAAATCATCCGAGTCCTGTTT 88911

Qy         158 ---SerProSerThrGlnSerSer------GlnGly------------LysAsnAlaAsp 170
                   |||||||||    :::      :::|||            |||:::   |||
Db       88912 CCTGGGAACAGCACCCAGTACAACATCCTGGAAGGGTTGGAAAAGCTCAAACATCTAGAC 88971

Qy         171 ArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSer 190
                     |||   :::   |||:::              :::         |||
Db       88972 TTTCTGAAGCAGCCACTGGCCACCCAAAAGGATCTCACTGGCCAGGTGCCCACTCCTGTG 89031

Qy         191 GlyLysAsnGlyAlaAlaGluAlaAsnThrAspSer------ProMetLysAspLeuGln 208
                  |||           :::         ||||||         :::|||
Db       89032 GTGAAACAAACAAAACTGAAACAGAGGGCTGATAGCCGAGAAAGTCTGAAGCCAGCCGCA 89091

Qy         209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
                  |||      :::      |||              |||
Db       89092 AAACCACTTCCTAGCAAATCCGTGCGCAAGGAGTCAAAAGAA----------------- 89133

Qy         229 ValGlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAsp 248
                        |||   |||::::::|||          :::||||||       :::
Db       89134 --------------GAAACCCCTGAGGTCACA---------AAAGTGAATCACGTGGAA 89169

Qy         249 GlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsn 268
               :::||||||                 :::|||::::::|||:::::::::
Db       89170 AAGCCACCCAAAGTTGAAAGCAAAGAAAAGGTAATGGTGAAAAAAGACAAGCCAGTAAAA 89229

Qy         269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
                  :::|||:::   ::::::      ::::::            ::::::|||:::|||
Db       89230 ACAGAGACCAAACCTTCAGTGACTGAAAAGGAGGTTCCCAGCAAAGAAGAGCCATCTCCA 89289

Qy         289 MetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLys 308
                  :::            :::   :::||||||||||:::                |||   |||
Db       89290 GTG------------AAAGCCGAGGTGGCTGAGAAG------CAAGCCACAGATGTCAAA 89331

Qy         309 ProValSerGly--------------GlnLysCysGluGlnIleCysAsnGluProCys 323
                  |||   :::                :::|||      ::::::      |||
Db       89332 CCCAAAGCTGCCAAGGAGAAGACGGTGAAAAAGGAAACAAAGGTAAAGCCTGAAGACAAG 89391

Qy         324 GluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMet 343
                  :::|||      |||      :::      ::::||   ||||||      :::
Db       89392 AAAGAGGAGAAAGAAAAGCCAAAGAAAGAAGTGGCTAAAAAGGAGGACAAAACACCTATC 89451

Qy         344 LysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspVal--------------- 358
                  ||||||::::::      ||||||          :::   :::|||
Db       89452 AAGAAGGAGGAAAAACCAAAAAAGGAAGAGGTGAAAAAAGAAGTCAAAAAAGAGATCAAG 89511

Qy         359 ---SerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIle 377
                     :::   |||      :::::::::   ||||||                |||:::
Db       89512 AAAGAAGAGAAAAAAGAACCCAAGAAAGAGGTTAAGAAAGAAACACCGCCAAAGGAAGTC 89571

Qy         378 IleAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAla 397
                     :::|||:::                  ||||||   :::|||
Db       89572 ------AAGAAGGAAGTTAAGAAGGAAGAGAAGAAGGAAGTGAAAAAGGAAGAAAAGGAA 89625

Qy         398 Asp--------------------ProCysGluAspAspArgSerThrIleProVal--- 409
                                      |||   :::      :::|||:::   |||:::
Db       89626 CCCAAAAAAGAAATTAAGAAGCTCCCTAAAGACGCAAAGAAATCATCTACTCCTCTGTCT 89685

Qy         410 ------------ProMetGluValSerMetAspIleProValSerAsnHisThrValGly 425
```


Db 89686 GAAGCAAAAAAACCAGCTGCTTTAAAACCAAAAGTACCCAAGAAGGAAGAGTCTGTCAAG 89745
Qy 426 GluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAsp 445
Db 89746 AAAGATTCTGTTGCTGCCGGAAAGCCAAAGGAGAAGGGGAAAATAAAAGTCATTAAGAAG 89805
Qy 446 AspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHis 465
Db 89806 GAAGGCAAGGCCGCAGAG------GCTGTCGCTGCAGCTGTCGGCACTGGAGCCACCACA 89859
Qy 466 HisThrValAlaHisProAlaGlyAsnLeuSer---------AsnLysLysValThrPro 482
Db 89860 GCAGCTGTCATGGCGGCAGCTGGAATAGCAGCCATTGGCCCTGCCAAAGAACTCGAAGCT 89919
Qy 483 ThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMet 502
Db 89920 GAGAGGTCCCTTATGTCATCTCCTGAGGATCTAACCAAGGACTTTGAAGAGTTAAAGGCT 89979
Qy 503 HisLysThrAspValCysGlnHisValSerGlu-------------------------- 513
Db 89980 GAAGAGGTCGATGTAACAAAGGACATCAAGCCTCAGCTGGAGCTAATCGAAGACGAAGAG 90039
Qy 514 ----------------------------------IleSerThrGlnArgCysSerSer 521
Db 90040 AAACTGAAGGAAACTGAGCCAGTCGAAGCCTACGTCATCCAGAAGGAGAGAGAAGTCACC 90099
Qy 522 LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyr 541
Db 90100 AAAGGTCCTGCCGAGTCCCCTGATGAGGGAATCACT-----------ACCACTGAAGGG 90147
Qy 542 GlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeu---------SerAlaGlu 558
Db 90148 GAGGGCGAATGTGAACAGACACCTGAGGAGCTGGAGCCCGTCGAGAAGCAGGGAGTAGAC 90207
Qy 559 AspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysVal------ 576
Db 90208 GACATTGAAAAATTTGAAGATGAAGGAGCCGGTTTTGAAGAATCTTCAGAGACTGGAGAC 90267
Qy 577 --SerProAlaGluHisAsp 582
Db 90268 TATGAAGAGAAGGCAGAAACTGAGGAGGCTGAGGAGCCAGAAGAGGATGGGGAGGAACAC 90327
Qy 583 IleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLys 602
Db 90328 GTATGTGTGAGCGCCTCCAAGCACAGCCCCACTGAGGATGAGGAAAGTGCCAAGGCGGAG 90387
Qy 603 LeuGluVal---ThrArgGluLysGlnThrMetIle------------AspAspIlePro 617
Db 90388 GCTGATGCATACATCAGGGAGAAGAGGGAGTCTGTGGCCAGTGGGGATGACCGAGCCGAA 90447
Qy 618 MetAspIleValGluLeuLeuAlaLys---------AsnGlnHisGluArgGlnLeuMet 634
Db 90448 GAAGACATGGATGAGGCCATTGAGAAAGGAGAGGCTGAACAATCTGAAGAGGAGGCTGAT 90507
Qy 635 ThrGluThrAspCysSerAspIleAsnArgIleGln---------SerLysThrThrAla 651
Db 90508 GAGGAGGACAAAGCTGAAGATGCCAGAGAGGAGGAATATGAGCCGGAAAAAATGGAAGCT 90567
Qy 652 AspAspAspCysValIleVal-- 658
Db 90568 GAAGACTATGTGATGGCTGTGGTCGACAAGGCTGCAGAGGCTGGTGGTGCCGAGGAGCAG 90627
Qy 659 ----------------------------------AlaAlaLysAspGlySerAspTyr 666
Db 90628 TATGGATTCCTCACCACACCAACCAAGCAACTAGGAGCCCAGTCTCCTGGCCGAGAACCT 90687
Qy 667 AlaSerSerValPheAspThr------------AsnSerGlnGlnLysSerLeuAlaSer 682
Db 90688 GCATCTTCAATTCATGATGAGACTTTACCTGGAGGCTCAGAGAGCGAGGCCACCGCTTCT 90747
Qy 683 GlnSerThrGlnLysGlu----------------------------------LeuGln 690
Db 90748 GATGAGGAGAATCGAGAAGACCAGCCTGAGGAATTCACTGCCACCTCTGGCTACACTCAG 90807
Qy 691 GlyHisLeuAlaLeuThrThrGlnGluSerPro---------------HisProGlnAsn 705
Db 90808 TCTACTATTGAGATATCCAGTGAGCCCACCCCCATGGATGAGATGTCTACCCCTCGAGAC 90867
Qy 706 PheGlnSer------ThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThr 723
Db 90868 GTGATGAGTGATGAGACCAACAATGAAGAGACGGAGTCCCCTTCTCAGGAATTCGTAAAT 90927
Qy 724 IleAlaAla---SerSerProLeuPheSerHisHis------------------------ 734
Db 90928 ATCACCAAATATGAATCTTCATTGTATTCTCAGGAATACTCTAAACCTGCTGATGTTACA 90987
Qy 735 --AspAspGlnTyrIle 739
Db 90988 CCGCTCAACGGATTTTCTGAAGGATCAAAAACAGATGCCACTGATGGCAAGGATTACAAT 91047
Qy 740 AlaGluAlaProThr--------------------------GluHisTrpGlyArgLys 750
Db 91048 GCTTCAGCCTCTACCATATCACCACCCTCTTCCATGGAGGAAGACAAATTCAGCAGATCT 91107
Qy 751 AspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAla 770
Db 91108 GCTTTACGTGATGCTTACTGCTCTGAAGTGAAAGCCAGCACCACT--------------- 91152
Qy 771 ThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
Db 91153 ---------TTGGACATCAAAGATAGCATCTCAGCTGTT------TCAAGTGAAAAGGTC 91197
Qy 791 MetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArg--- 809
Db 91198 AGCCCATCGAAGAGCCCGTCCCTGAGTCCATCTCCACCA---TCACCCTTAGAAAAGACC 91254
Qy 810 ---TyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAla 828
Db 91255 CCCCTGGGTGAACGTAGTGTGAACTTCTCTCTGACGCCCAATGAGATTAAAGTCTCTGCA 91314
Qy 829 ThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPro 848
Db 91315 GAGGCAGAAGTAGCCCCGGTGTCTCCT------GAGGTGACCCAAGAAGTAGTTGAAGAA 91368
Qy 849 LysGluSerMetProAlaThrHisLeuArgMetMetAspProSerThrLeuAlaSer 868
Db 91369 CATTGTGCTAGTCCTGAGGACAAGACTCTGGAAGTGGTGTCACCA--------------- 91413
Qy 869 PheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGln 888
Db 91414 ------------------------------------TCTCAGTCCGTGACTGGCAGT 91434
Qy 889 TyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGly 908
Db 91435 GCTGGTCACACACCTTACTATCAATCTCCTACTGACGAGAAATCCAGTCATCTCCCTACA 91494
Qy 909 Lys---------------------IleProLeuThrPheGlu------------------ 915
Db 91495 GAAGTCATTGAAAAACCACCAGCAGTTCCAGTGAGTTTTGAATTCAGTGATGCCAAAGAT 91554
Qy 916 --AspLeuSerArgHisGln 921
Db 91555 GAGAATGAAAGGGCTTCAGTAAGCCCCATGGATGAGCCCGTGCCTGACTCAGAGTCTCCT 91614
Qy 922 LeuHisAspLeuHisArgProLeuArgProHisProArgValGly---------ValLeu 938
Db 91615 ATTGAAAAAGTTTTGTCTCCTTTACGCAGCCCGCCCCTCATTGGATCCGAGTCTGCTTAT 91674
Qy 939 GlySerLeuLeuGlnLysGlu----------------------IleAlaAsnTrpSer 950
Db 91675 GAAAGTTTTCTAAGTGCTGATGACAAGGCTTCTGGCAGAGGTGCCGAAAGTCCTTTTGAA 91734
Qy 951 GluAsnCysGlyThrGln---SerGlyTyrLysLeuGlyValSerThrGlyIleThrSer 969
Db 91735 GAAAAGAGTGGAAAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTTCTGAAATGACTTCT 91794

```
Qy        970 HisGlnMetAsnArg-------------------------------------LysGlu 976
Db      91795 ACTAGTCTTTACCAAGACAAACAGGAAGGGAAAAGCACAGACTTTGCACCAATAAAAGAA 91854

Qy        977 HisPhe--------------------------GluAlaLeuAsnSerGlyMetPheSer 987
Db      91855 GACTTTGGCCAAGAAAAGAAAACTGATGATGTTGAAGCCATGAGTTCTCAACCAGCACTG 91914

Qy        988 AlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSer------------AlaAsp 1003
Db      91915 GCT---CTGGATGAAAGGAAATTAGGAGATGTTTCTCCCACACAAATAGATGTCAGTCAG 91971

Qy       1004 PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGly------LysGlyLys 1021
Db      91972 TTTGGATCTTTTAAAGAAGACACTAAGATGTCCATTTCTGAAGGTACTGTCTCAGACAAG 92031

Qy       1022 MetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
Db      92032 TCAGCTACTCCTGTTGATGAGGGCGTAGCAGAAGACACGTACTCTCATATGGAGGGTGTG 92091

Qy       1042 AlaAspPheThrThrIleSer 1048
Db      92092 GCCTCAGTGTCCACAGCCTCA 92112

RESULT 13
US-09-949-016-13530
; Sequence 13530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13530
;   LENGTH: 145287
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13530

Alignment Scores:
Pred. No.:              1.6e-05       Length:         145287
Score:                  196.50        Matches:        207
Percent Similarity:     34.3%         Conservative:   164
Best Local Similarity:  19.1%         Mismatches:     370
Query Match:            3.6%          Indels:         341
DB:                     3             Gaps:           49

US-09-828-068-2 (1-1057) x US-09-949-016-13530 (1-145287)

Qy         28 ValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaAlaGluAlaGly 47
Db     126105 GTTACACCCAAAACCACAGATGTAAGTACACCAAAACCAGCTGTGATTCATGAATGTGCA 126164

Qy         48 ValAspGluProAlaGlnHis-----------------------GlnCysGluHisPhe 59
Db     126165 GAGGAGGATGATTCAGAAAACGGGGAGAAAAAGAGGTTCACACCTGAAGAGGAGATGTTT 126224

Qy         60 SerIleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSer 79
Db     126225 AAAATGGTAACCAAAATCAAAATGTTTGATGAACTTGAACAAGAA---GCAAAGCAGAAA 126281

Qy         80 ArgIlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSer 99

Db     126282 AGGGACTACAAAAAAGAACCCAAACAAGAAGAATCTTCTTCATCTTCTGACCCAGATGCT 126341

Qy        100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
Db     126342 --------------------GACTGTTCAGTAGATGTGGATGAACCAAAACATACAGGC 126380

Qy        120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
Db     126381 AGTGGGGAGGATGAAAGTGGTGTCCCT-------------------------------- 126407

Qy        140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerPro 159
Db     126408 ----------------------GTGTTAGTAACTTCGGAGAGCAGGAAGGTGTCTTCC 126443

Qy        160 SerThrGlnSerSerGln---------------GlyLysAsnAlaAspArgSerThrLeu 174
Db     126444 TCCTCAGAAAGTGAACCTGAGTTGGCACAGCTTAAAAAAGGTGCTGACTCAGGCCTTTTA 126503

Qy        175 ProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGly 194
Db     126504 CCAGAACCAGTG---------------------------------------------- 126515

Qy        195 AlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyr 214
Db     126516 ---ATTCGAGTACAACCTCCTTCTCCACTTCCA-------------------------- 126545

Qy        215 AspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGlu 234
Db     126546 ----------------------------TCAAGCATGGACTCCAATTCCAGTCCAGAA 126575

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrPro 254
Db     126576 GAAGTACAATTC----------------------------CAGCCTGTCGTTTCC--- 126602

Qy        255 LysLeuSerGluValValLeuLysArgAsnGluAsp---------GluAsnGlyLysThr 271
Db     126603 ------AAACAATATACTTTCAAGATGAATGAAGATACTCAGGAAGAGCCAGGCAAATCA 126656

Qy        272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
Db     126657 GAAGAAGAAAAAGATTCTGAATCC---CATTTAGCTGAAGACCGTCATGCTGTTTCCACT 126713

Qy        292 LysGluArgAspGlnValAlaGluGlnCysAsnLeu---ThrLysAspProLysProVal 310
Db     126714 GAGGCTGAAGACAGGTCTTATGATAAGCTAAACAGAGACACTGATCAGCCAAAAATCTGT 126773

Qy        311 SerGlyGlnLysCysGluGlnIleCys---------------------------------- 319
Db     126774 GATGGCCATGGATGTGAGGCCATGAGTCCTAGCAGCTCAGCTGCTCCTGTCTCTTCAGGT 126833

Qy        320 -----------------AsnGluProCysGluGluValValLeuLysArgSerSerLys 333
Db     126834 CTACAGAGTCCGACTGGTGATGATGTTGATGAACAGCCAGTCATCTATAAAGAATCATTA 126893

Qy        334 SerLysArgLysThrAspLysLysLeuMetLysLysGlnGln------HisSerLysLys 351
Db     126894 GCTCTCCAAGGCACTCATGAAAAAGACACAGAGGGAGAAGAGCTTGATGTTTCTAGAGCA 126953

Qy        352 ArgThrAlaGlnAlaAspVal-----------------SerAspAlaLysLeuCysArg 365
Db     126954 GAATCTCCACAAGCAGATTGCCCCAGTGAAAGCTTTTCATCTTCATCCTCTTTGCCTCAT 127013

Qy        366 ArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla------------Asn 381
Db     127014 TGTTTGGTATCTGAAGGAAAAGAATTAGATGAAGACATATCTGCCACATCTTCTATTCAA 127073

Qy        382 GlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGlu 401
Db     127074 AAAACAGAGGTCACAAAAACTGATGAAACATTT---GAGAACTTACCAAAGGACTGCCCC 127130

Qy        402 AspAspArgSerThrIleProValProMetGlu---ValSerMetAspIleProValSer 420
```





RESULT 14
US-09-949-016-13531
; Sequence 13531, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13531
;   LENGTH: 145287
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13531

Alignment Scores:
Pred. No.:                1.6e-05        Length:          145287
Score:                    196.50         Matches:         207
Percent Similarity:       34.3%          Conservative:    164
Best Local Similarity:    19.1%          Mismatches:      370
Query Match:              3.6%           Indels:          341
DB:                       3              Gaps:            49

US-09-828-068-2 (1-1057) x US-09-949-016-13531 (1-145287)

Qy         28 ValAlaProValLeuGluLeuThrAlaThrProArgGlnAspAlaAlaAlaGluAlaGly 47
              |||   |||                  :::|||||||:::                 |||
Db     126105 GTTACACCCAAAACCACAGATGTAAGTACACCAAAACCAGCTGTGATTCATGAATGTGCA 126164

Qy         48 ValAspGluProAlaGlnHis-----------------------GlnCysGluHisPhe 59
                 ::::::   :::::::::                     :::   |||   |||
Db     126165 GAGGAGGATGATTCAGAAAACGGGGAGAAAAAGAGGTTCACACCTGAAGAGGAGATGTTT 126224

Qy         60 SerIleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSer 79
              :::           :::   :::      :::   :::   :::
Db     126225 AAAATGGTAACCAAAATCAAAATGTTTGATGAACTTGAACAAGAA---GCAAAGCAGAAA 126281

Qy         80 ArgIlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSer 99
              |||   :::      :::   |||   :::|||      :::|||||   |||   :::
Db     126282 AGGGACTACAAAAAAGAACCCAAACAAGAAGAATCTTCTTCATCTTCTGACCCAGATGCT 126341

Qy        100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
                                  |||||||||      :::|||:::   |||   :::
Db     126342 --------------------GACTGTTCAGTAGATGTGGATGAACCAAAACATACAGGC 126380

Qy        120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
              :::|||                 :::|||
Db     126381 AGTGGGGAGGATGAAAGTGGTGTCCCT---------------------------------- 126407

Qy        140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerPro 159
                                     |||      :::      |||:::|||||||||
Db     126408 ----------------------GTGTTAGTAACTTCGGAGAGCAGGAAGGTGTCTTCC 126443

Qy        160 SerThrGlnSerSerGln--------------GlyLysAsnAlaAspArgSerThrLeu 174
              |||::::::|||                  |||   ||||||           |||
Db     126444 TCCTCAGAAAGTGAACCTGAGTTGGCACAGCTTAAAAAAGGTGCTGACTCAGGCCTTTTA 126503

Qy        175 ProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGly 194
              |||:::   |||
Db     126504 CCAGAACCAGTG---------------------------------------------- 126515

Qy        195 AlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyr 214
                                  ||||||:::
Db     126516 ---ATTCGAGTACAACCTCCTTCTCCACTTCCA------------------------- 126545

Qy        215 AspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGlu 234
                                        :::|||:::|||      :::   ||||||
Db     126546 ----------------------------TCAAGCATGGACTCCAATTCCAGTCCAGAA 126575

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrPro 254
                 |||                                |||||      :::
Db     126576 GAAGTACAATTC----------------------------CAGCCTGTCGTTTCC--- 126602

Qy        255 LysLeuSerGluValValLeuLysArgAsnGluAsp---------GluAsnGlyLysThr 271
                    :::         |||   |||||||||          |||   ||||||:::

Db     126603 ------AAACAATATACTTTCAAGATGAATGAAGATACTCAGGAAGAGCCAGGCAAATCA 126656

Qy        272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
              ||||||          :::|||      :::|||   :::|||   :::   :::|||
Db     126657 GAAGAAGAAAAAGATTCTGAATCC---CATTTAGCTGAAGACCGTCATGCTGTTTCCACT 126713

Qy        292 LysGluArgAspGlnValAlaGluGlnCysAsnLeu---ThrLysAspProLysProVal 310
              :::      |||:::      ::::::   |||      |||      ||||||
Db     126714 GAGGCTGAAGACAGGTCTTATGATAAGCTAAACAGAGACACTGATCAGCCAAAAATCTGT 126773

Qy        311 SerGlyGlnLysCysGluGlnIleCys------------------------------- 319
                |||      ||||||   :::
Db     126774 GATGGCCATGGATGTGAGGCCATGAGTCCTAGCAGCTCAGCTGCTCCTGTCTCTTCAGGT 126833

Qy        320 -----------------AsnGluProCysGluGluValValLeuLysArgSerSerLys 333
                              ::::::      |||:::   |||:::   :::   |||
Db     126834 CTACAGAGTCCGACTGGTGATGATGTTGATGAACAGCCAGTCATCTATAAAGAATCATTA 126893

Qy        334 SerLysArgLysThrAspLysLysLeuMetLysLysGlnGln------HisSerLysLys 351
              :::   :::   |||   :::|||      :::   ::::::         |||:::
Db     126894 GCTCTCCAAGGCACTCATGAAAAAGACACAGAGGGAGAAGAGCTTGATGTTTCTAGAGCA 126953

Qy        352 ArgThrAlaGlnAlaAspVal-----------------SerAspAlaLysLeuCysArg 365
                 :::   |||||||||                   |||   :::   |||
Db     126954 GAATCTCCACAAGCAGATTGCCCCAGTGAAAGCTTTTCATCTTCATCCTCTTTGCCTCAT 127013

Qy        366 ArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla------------Asn 381
                          :::   :::   |||   |||   |||:::|||
Db     127014 TGTTTGGTATCTGAAGGAAAAGAATTAGATGAAGACATATCTGCCACATCTTCTATTCAA 127073

Qy        382 GlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGlu 401
              :::   |||   :::::::::||||||          ||||||                |||
Db     127074 AAAACAGAGGTCACAAAAACTGATGAAACATTT---GAGAACTTACCAAAGGACTGCCCC 127130

Qy        402 AspAspArgSerThrIleProValProMetGlu---ValSerMetAspIleProValSer 420
                  |||:::|||             :::      |||||||||:::|||||||||
Db     127131 TCTCAAGACTCATCCATTACTACTCAAACAGATAGATTTTCCATGGATGTTCCCGTGTCT 127190

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
              :::      :::   |||:::
Db     127191 GAC------CTAGCTGAGAAT---------------------------------------- 127205

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
                 ::::::   |||
Db     127206 GATGAAATCTATGATCCA---------------------------------------- 127223

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
                                     :::      |||      |||:::
Db     127224 ------------------CAAATCACTAGCCCTTATGAAAATGTC-------------- 127250

Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
                 |||:::   |||            :::|||:::   |||:::        |||
Db     127251 ---CCTTCCCAATCTTTTTTCTCTAGTGAAGAAAGCAAAACCCAAACA------GATGCA 127301

Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
              |||   |||   |||            |||   ||||||:::   :::         :::
Db     127302 AAT---CACACCACAAGTTTT------CACTCTTCTGAAGTGTATTCTGTTACCATCACA 127352

Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
              |||      :::         ::::::                  |||::::::   |||
Db     127353 TCCCCTGTTGAAGACGTTGTAGTGGCAAGC-----------TCCTCTAGTGGAACTGTT 127400

Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
                 ||||||          ||||||:::|||   :::   |||   ::::::
Db     127401 TTAAGCAAAGAATCTAATTTTGAGGGCCAGGACATAAAAATGGAATCCCAACAGGAAAGT 127460

Qy        561 CysGlnMetGluThrGluAsnSerValLeuSerHisSer--------------------- 573
                    |||   ::::::      :::|||   |||
Db     127461 ACCTTGTGGGAAATGCAATCAGACAGTGTCTCTTCATCTTTCGAGCCTACTATGTCCGCT 127520

Qy        574 -------------------------AlaLysValSerProAlaGluHisAspIle--- 583
                                        :::|||||||         :::   |||:::
Db     127521 ACAACAACAGTTGTTGGTGAACAAATAAGCAAAGTCATCATCACAAAAACTGATGTGGAT 127580
```

```
Qy       584 --------------GlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLys 598
                           :::|||   |||                          |||:::
Db    127581 TCTGATTCTTGGAGTGAAATTCGGGAAGACGATGAAGCCTTTGAGGCTCGTGTGAAAGAG 127640

Qy       599 LysLysGlnLys---------LeuGluValThrArgGluLysGln---ThrMetIleAsp 614
             ::::::||||||         |||  |||  |||:::  |||  |||      |||
Db    127641 GAAGAACAAAAGATATTTGGTTTGATGGTAGACAGACAATCACAGGGTACCACCCCTGAC 127700

Qy       615 AspIlePro-------------------------------------------------- 617
                 |||
Db    127701 ACCACTCCTGCTAGGACCCCAACTGAAGAGGGGACCCCAACAAGTGAGCAAAACCCATTT 127760

Qy       618 ----------------------------------------------MetAspIleVal--- 621
                                                          :::|||:::
Db    127761 CTGTTTCAGGAAGGAAAATTGTTTGAAATGACCCGAAGTGGTGCCATTGATATGACCAAA 127820

Qy       622 ------------------------------------------------------Glu 622
                                                                    |||
Db    127821 AGGTCCTATGCAGATGAAAGTTTTCACTTTTTCCAAATTGGTCAAGAATCCAGGGAAGAG 127880

Qy       623 LeuLeuAlaLysAsnGlnHisGlu--------------ArgGlnLeuMetThrGluThr 637
               |||:::::::::      |||                |||      ||||||
Db    127881 ACTCTCTCTGAAGATGTGAAAGAAGGGGCTACTGGGGCTGATCCCCTACCGCTGGAGACA 127940

Qy       638 AspCysSerAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIle 657
                    :::        :::||||||   |||  ||||||:::      :::
Db    127941 TCAGCTGAATCACTAGCACTTTCAGAATCAAAAGAAACAGTGGATGATGAGGCAGACTTA 128000

Qy       658 Val---------AlaAlaLysAspGlySerAspTyrAlaSerSerValPheAspThrAsn 674
             :::            ::::::::      :::    :::|||              |||
Db    128001 CTTCCAGATGACGTGAGTGAGGAAGTAGAGGAAATACCTGCTTCGGATGCTCAACTTAAC 128060

Qy       675 SerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAla 694
             ||||||                |||   ||||||:::          :::      ::::::
Db    128061 TCCCAAATGGGG------ATTTCAGCCTCCACTGAAACACCTACAAAAGAAGCTGTTAGT 128114

Qy       695 LeuThrThrGlnGluSerProHisProGln---------------------------- 704
             :::   |||::::::   |||      |||
Db    128115 GTAGGGACCAAGGACCTCCCCACCGTGCAAACGGGTGATATACCTCCTCTCTCTGGTGTA 128174

Qy       705 ---------AsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMet 721
                    :::        :::   |||              :::::::::   :::
Db    128175 AAGCAGATATCCTGCCCCGACTCTTCTGAACCAGCTGTACAAGTCCAGTTA---GATTTT 128231

Qy       722 ValThrIleAlaAlaSer-----------------------SerProLeuPheSerHis 733
               |||:::      |||                      ||||||      |||
Db    128232 TCCACACTCACCAGGTCTGTTTATTCAGATAGGGGTGATGATTCTCCCGATTCTTCCCCA 128291

Qy       734 HisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLys 753
               :::   :::   :::   |||   ||||||                  :::
Db    128292 GAAGAACAGAAATCAGTAATCGAGATTCCTACT---------GCACCCATGGAGAATGTG 128342

Qy       754 LysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerProAlaAlaThr------ 771
                |||    :::  |||     |||     |||   :::|||
Db    128343 CCTTTTACTGAAAGCAAATCCAAAATTCCTGTAAGGACTATGCCCACTTCCACCCCAGCA 128402

Qy       772 ---CysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerThrHisVal 790
                  |||::::::         :::             |||:::|||
Db    128403 CCTCCATCTGCAGAGTATGAGAGTTCAGTTTCTGAAGATTTTCTATCCAGTGTA------ 128456

Qy       791 MetGlySerSerSerAsnTyrAla-----------------SerArgGlnProValIle 804
                           |||   |||                  |||:::   ||||||
Db    128457 ------GATGAGGAAAATAAGGCGGATGAAGCAAAACCAAAGTCCAAACTCCCTGTCAAA 128510

Qy       805 AlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPhePro 824
             ||||||   |||     |||:::   ::::::   :::         :::
Db    128511 GTACCCCTCCAAAGA---GTTGAACAGCAGCTCTCAGATCTAGACACC---TCTGTCCAG 128564

Qy       825 SerThrIleAlaThrMet---------GluAlaSerLysLeuCysAspArgArgAsnAla 841
               |||:::|||              ||||||          |||   |||:::
Db    128565 AAGACAGTGGCTCCTCAGGGACAGGACATGGCAAGCATCGCACCAGATAATAGAAGCAAA 128624

Qy       842 GlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMet 861
                  :::
Db    128625 TCTGAA----------------------------------------------------TCT 128633

Qy       862 AspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGlu 881
                |||   |||:::|||   |||                        :::|||:::      |||
Db    128634 GATGCTAGTTCTTTGGATTCAAAGACCAAATGCCCAGTAAAAACCCGAAGTTACACTGAG 128693

Qy       882 SerGlnLeuHisAsnSerGlnTyrAlaHis------------AsnGlnTyrLysGlySer 897
             ::::::      :::   :::   |||                    :::|||:::
Db    128694 ACAGAAACAGAGAGCAGAGAGAGGGCCGAGGAACTTGAGTTAGAATCAGAAGAAGGGGCC 128753

Qy       898 Thr-----------------------------------SerThrSerTyrGlySerAsn 905
             |||                                  |||   |||      |||:::
Db    128754 ACAAGACCAAAGATACTTACATCCCGATTGCCAGTTAAGAGCAGAAGCACTACATCTTCC 128813

Qy       906 LeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeu 925
                |||              |||   |||      |||:::            ||||||
Db    128814 TGCAGGGGGGGCACGAGCCCCACAAAAGAA------AGTAAGGAGCATTTCTTTGACCTT 128867

Qy       926 HisArg 927
             :::|||
Db    128868 TACAGA 128873
```

RESULT 15
US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
; GENERAL INFORMATION:
; APPLICANT: Iandolo, John J.
; APPLICANT: Crupper, Scott S.
; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,999
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/710,561
; FILING DATE: 19-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 25043-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 816/474-9050
; TELEFAX: 816/474-9057
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6755 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; STRAIN: UT0007
US-08-931-999-4

Alignment Scores:

Pred. No.:	8.24e-08	Length:	6755
Score:	196.00	Matches:	160
Percent Similarity:	31.1%	Conservative:	101
Best Local Similarity:	19.0%	Mismatches:	403
Query Match:	3.5%	Indels:	176
DB:	3	Gaps:	27

US-09-828-068-2 (1-1057) x US-08-931-999-4 (1-6755)

```
Qy        22 ArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGlnAsp  41
             ||||||           |||                          ||||||||
Db      2787 AGGGGGGCAAAGAACGGCCAAACGCCCACGCAAACAAAAAACCAAGGGGAAAGACAGGAC 2846

Qy        42 AlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSerIle  61
                 |||:::::::|||      :::         |||:::
Db      2847 AGGCGAGCAGACAGCGGCAGGCACAAACAACAAACACACGAAGCGACGGCAACGAAAAAG 2906

Qy        62 ArgGlyTyrValAlaLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle  81
             :::                 |||||||||   |||:::                 :::
Db      2907 AAA-----------------CAGAAGAAGAAACCCGAG------GAGAACGAAAAGCAG 2942

Qy        82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerValAla  101
                   |||::::::      |||   :::
Db      2943 ACAAAGCAGCAGGAGAGAAAAAGAAGCACAGACACACCAGAAAGCCCAGCAAAAAACA 3002

Qy       102 LysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsnGly  121
             |||   :::|||   |||   |||:::                |||       :::
Db      3003 AAAAGAAAAAGAGGGGACACAAGCAGGGGCAGCGGGACAGGAAAACACAAGGAAACAAGA 3062

Qy       122 ThrAlaProArgThrLeuProAla---------------LysGlnAsnGlyThrSerAsp  136
                 ||||||                          ||||||      |||   |||
Db      3063 CCCGAACCCAGGGCAGCAAAGGGGGGAGAGAAAACAAAAAAACAGAAAAAGACACAAGAC 3122

Qy       137 GlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLys  156
                        |||:::                          |||:::   :::
Db      3123 AACCACAAGAAGCAACCCAAAAGAACAAGG---------AAAAGAAGGGGGAACACAAGA 3173

Qy       157 ValSerProSerThrGlnSerSerGlnGlyLysAsnAla---------------------  169
                   |||   :::||||||   |||:::|||
Db      3174 AACGAAAAACAAACGGCCAACAGCCAAAACAAAAGCGCCAGGGAAAGACAGAGCGAAAGG 3233

Qy       170 ------AspArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsn  187
                   :::|||       |||       :::       |||       :::
Db      3234 GGCAAGCAGGAGGCAACGACAAAGAAACAAAGCAAAAGCCGCAACCCAAAAAGAGGAGAA 3293

Qy       188 AlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLys------  205
             ||||||:::   |||      |||   :::   ||||||          |||
Db      3294 GCACCCACG---AAAGGCGCAGCACAGAAGACGAACACGAGAAAAAGCACCAAAACGAAG 3350

Qy       206 ----------------------------AspLeuGlnGlyProAlaGlnAsnTyrAsp  215
                                          :::|||   |||         :::
Db      3351 CCACGAGCACGAAAGCAGAAAAAAGGAGGACAGGAGAAAGGAAGGGCGCACAAAGAAGAA 3410

Qy       216 ValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluVal  235
                       ::::::          :::
Db      3411 AAAAAAGGAAAGAAAACAAAGCGAAAAAAAGGAGGGGAG-------------------- 3449

Qy       236 ProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrPro---  254
                        |||          :::|||       ||||||
Db      3450 --------------------GAGAAAAGAAAAAGAGAACAA------AAAACACCGAGC 3482

Qy       255 --------------------LysLeuSerGluValValLeuLysArgAsnGluAspGlu  267
                             |||   ::::::         :::|||   :::::::|||
Db      3483 AGAAGAAAAAAAAAAGGGAAAAAAAAAACAAAAAAAAAAAACAGAAGACAAGACGAGGAG 3542

Qy       268 AsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsn  287
                   |||       :::          :::|||          ||||||
Db      3543 AAAAAAAAAAAGAACAAGCAGACAACGAAGAAACAAGCGGAAAACACGAAA---AAACGA 3599

Qy       288 ProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspPro  307
                   |||:::::::   |||       |||          |||   |||
Db      3600 GAACAGAGAGAAAAAAAAAAAACCCAAGCACAACCGCAAGCGACCAAAAAAAAAACGCCG 3659
```

```
Qy       308 LysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValVal  327
             ||||||                                  :::|||   |||
Db      3660 AAACCAAACAGG----------------------AAAAAACCAAAAGAGAGAAAGAAA 3695

Qy       328 LeuLysArgSerSerLysSerLysArgLysThrAsp------LysLysLeuMetLysLys  345
               |||:::   ||||||   :::|||||||||:::       :::|||      ||||||
Db      3696 AAAAAAAAACGGAGCAAAAAGCGCAGAAAAACGGAGGAAACAAGAAAAAAACCAAAAAAA 3755

Qy       346 GlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArg  365
               :::      |||  ||||||                    |||  :::      :::
Db      3756 GGGAAAAAGCAAAAAACGCGGACAGACCACAAAAAAAAAAAGGGACAGGAGAAAAGCAAAA 3815

Qy       366 ArgLysProLysLysValArgLeuLeuSerGluIleIle---AsnAlaAsnGln------  382
             ||||||  |||                |||             ||||||:::
Db      3816 GAAAAACCAAACAAAAGAGAAGCACGAAAAGAAAAAACAGGAAGAGCGAACAAAAAACCG 3875

Qy       383 ---ValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGlu  401
                ||:::         :::       |||:::       ||||||   :::
Db      3876 GCACAAGAAGAAGAGGAGAAGGAGGCAACGAAACGAAAACCAAAAAGAGACCCAAAGCAA 3935

Qy       402 AspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSerAsn  421
                :::                                              :::
Db      3936 CAAAGGCAGCGCAACGCCAGGGCACAGAGCAGAGAGAAAAACAAAACGAAACCAAACGAG 3995

Qy       422 HisThrValGly--------------------GluAspGlyLeuLysSerSerLysAsn  434
             |||                            |||   |||   |||::::::
Db      3996 GAAACAGCCAGCCGAAAAAAAAAAAAACAAGAGAACAGCGGAAAAAAACCAAGCAGACAC 4055

Qy       435 LysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeu  454
             |||   |||::::::   ::::::
Db      4056 AAACCCAAAAAAAGAACAGCAGAA---------------------------------- 4079

Qy       455 AsnGly--------------------LysLysLysArgThrGlySerValHisHisThr  467
               |||                          :::|||:::|||
Db      4080 AGAGGGCAAAAACCCCACCCCAAAAAACAAAAAAGGCGCAAAAAAGGCACGAGCAGGAAA 4139

Qy       468 ValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGln  487
                           |||      :::   |||:::         :::
Db      4140 GCCAAAGAAAAGACCGGAACCAAAAACCAAAAACAACAGAAGGGAAGCAAGCAACCCACC 4199

Qy       488 HisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspVal  507
             |||:::|||            |||||||||       :::       ||||||
Db      4200 CACAACGACAGAAGGCAAAAGGAAAACGGGCGGAAAAACGACCAAAAAAAAACAAAAACA 4259

Qy       508 CysGlnHisValSerGluIleSerThrGlnArgCysSer---------------SerLys  522
                :::                          |||::::::   |||          |||
Db      4260 CACAAGAGAGAAAAAAGGCCACCGACAAAACAGCCGAGCAAAAACCCCAGAAAACAAAAA 4319

Qy       523 GlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys------  540
             ||||||   |||                |||   |||   |||          |||
Db      4320 GGAAAAAAAGCG------------AAAAAAAGACACAAAGCCAACAAAAGAAAGAAGAAA 4367

Qy       541 ---------TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAla  557
                           :::   ::::::|||   :::|||   :::       :::
Db      4368 AAGCCGGAAAAAAGAAGAAAAAAAAGCAAAAACAAGGAGAACCCCAACAAACAAGCAAAA 4427

Qy       558 GluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSer  577
                :::   |||   :::   |||                              |||
Db      4428 AAAACACAAGGACGACAGGAA----------------------------AAGCGAGAA 4457

Qy       578 ProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLys  597
             |||   |||                                          ::::::
Db      4458 CCCAGAGAA-------------------------------------AAAAGGGAA 4475

Qy       598 LysLysLysGlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIlePro  617
             |||   |||||||||   |||       :::   :::
Db      4476 AAAAACAAACAAAAAAAGGAAAGCAAAAACCAAAGCAGGAAG---------------- 4517

Qy       618 MetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThr  637
                                 |||   |||   |||   :::       :::|||
Db      4518 --------------------GCAAACAACACAAAAGAGGCGGAGACGGAAGAAAAAACA 4556
```

```
Qy        638 AspCysSerAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIle 657
              :::       :::|||   |||      ::::::
Db       4557 CCCCGAAAAAAACAACAAAAAGACAACACAAAACCAACCCAAAGAGAAGAAAGAAACACA 4616

Qy        658 ValAlaAlaLysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln 677
              ::::::|||             :::|||      :::          |||
Db       4617 ACGAAAAAAGAAAACGGGGAAGGAAGAGGAAACAGCCAGAGAAACAGACAACCACAGAAC 4676

Qy        678 LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGly---------HisLeuAla 694
                  |||          |||      :::|||
Db       4677 GGGAGGGGACCAGAACAAAAAGAAGGCAAAACCAGAGAAGGGACAAACCAGCAAACAGGA 4736

Qy        695 LeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThr 714
              |||||||||           :::       |||   |||:::   :::    |||
Db       4737 ACAAGAACACAAGAA-----------AAAGAAAAACAAGAAACCGAAACGGAAAACACA 4784

Qy        715 HisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeuPheSerHisHis 734
                  |||                 :::
Db       4785 AGACAGAACAAAAACGAAAAACGACAAAGAAAAAGCGGCCGACGGACAAAAAAAGGAAGA 4844

Qy        735 AspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLys 754
              :::|||                 ||||||                |||      ::::::
Db       4845 GAGGACCCCAGGAAGGAGACACCCCCCACACCCAGAAAAAAAAACAAGACAAGGAGACGG 4904

Qy        755 LeuThrTrpGluGlnPheLysAlaThrThr---------------------------- 764
                  |||:::   |||||||||
Db       4905 AAAAACACAGAAAAAAGGAAAGCCACGGCAAAGGAAGGCAAAACGAAAAAAGAAAAGAAA 4964

Qy        765 --------------ArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGly 779
                      |||                 |||   |||   :::      ||||||
Db       4965 AGAACAGAAAAAAGGAGGAAGAAAAAAACCACAACGGGGGGGGGCAAGCAGGGACCAGGA 5024
```

Search completed: March 7, 2006, 16:48:38
Job time : 1010 secs

GenCore version 5.1.7

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2006, 14:08:04 ; Search time 1678 Seconds
(without alignments)
5209.018 Million cell updates/sec

Title: US-09-828-068-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTNCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/absa/ABSSWEB_spool/US09828068/runat_06032006_163200_25272/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=absa03p
-USER=US09828068_@CGN_1_1_1549_@runat_06032006_163200_25272 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5526	100.0	3896	3	US-09-828-068-1	Sequence 1, Appli
2	5526	100.0	4310	6	US-10-447-135-1	Sequence 1, Appli
3	5526	100.0	4640	7	US-10-437-963-40889	Sequence 40889, A
4	5191	93.9	9455	6	US-10-447-135-3	Sequence 3, Appli
5	491.5	8.9	871	7	US-10-425-114-4042	Sequence 4042, Ap
6	491.5	8.9	884	8	US-10-425-115-140777	Sequence 140777,
7	475	8.6	1039	8	US-10-425-115-176502	Sequence 176502,
8	317	5.7	3679	8	US-10-739-930-66	Sequence 66, Appl
9	241.5	4.4	7036	9	US-10-764-420-1635	Sequence 1635, Ap
10	223.5	4.0	5286	9	US-10-450-763-8960	Sequence 8960, Ap
11	223	4.0	2612	6	US-10-029-386-22571	Sequence 22571, A
12	223	4.0	10217	9	US-10-956-157-4592	Sequence 4592, Ap
13	223	4.0	10330	9	US-10-956-157-426	Sequence 426, App
14	223	4.0	10330	9	US-10-934-998-25	Sequence 25, Appl
15	223	4.0	10452	3	US-09-968-007A-490	Sequence 490, App
16	223	4.0	10452	9	US-10-843-641A-6960	Sequence 6960, Ap
17	223	4.0	10452	9	US-10-934-998-50	Sequence 50, Appl
18	223	4.0	11167	6	US-10-252-157-188	Sequence 188, App
19	222.5	4.0	279	3	US-09-294-093B-2759	Sequence 2759, Ap
20	221	4.0	5954	9	US-10-450-763-16395	Sequence 16395, A
c 21	221	4.0	6143	9	US-10-450-763-19647	Sequence 19647, A
22	221	4.0	6542	8	US-10-723-860-1889	Sequence 1889, Ap
23	221	4.0	7479	7	US-10-697-526-1	Sequence 1, Appli
c 24	220	4.0	3953	3	US-09-864-761-19041	Sequence 19041, A
c 25	220	4.0	3953	6	US-10-029-386-20602	Sequence 20602, A
26	220	4.0	4383	6	US-10-295-027-427	Sequence 427, App
27	219	4.0	6745	6	US-10-120-988-286	Sequence 286, App
28	214	3.9	4584	3	US-09-945-917-51	Sequence 51, Appl
29	214	3.9	4749	3	US-09-945-917-55	Sequence 55, Appl
30	214	3.9	5009	3	US-09-945-917-52	Sequence 52, Appl
31	214	3.9	5010	3	US-09-945-917-56	Sequence 56, Appl
32	214	3.9	5072	3	US-09-945-917-2	Sequence 2, Appli
33	214	3.9	5073	3	US-09-945-917-1	Sequence 1, Appli
34	214	3.9	9642	3	US-09-945-917-30	Sequence 30, Appl
35	214	3.9	10443	3	US-09-945-917-25	Sequence 25, Appl
36	214	3.9	13414	3	US-09-945-917-27	Sequence 27, Appl
37	213	3.9	2290	5	US-10-102-524-1741	Sequence 1741, Ap
38	213	3.9	2290	7	US-10-648-593-21	Sequence 21, Appl
39	213	3.9	8733	6	US-10-144-198-11	Sequence 11, Appl
40	212.5	3.8	6816	7	US-10-282-122A-7740	Sequence 7740, Ap
41	211.5	3.8	6783	8	US-10-470-048B-8	Sequence 8, Appli
42	211	3.8	7474	3	US-09-945-917-26	Sequence 26, Appl
43	209	3.8	6852	6	US-10-172-502-3	Sequence 3, Appli
44	209	3.8	6852	10	US-11-020-509-3	Sequence 3, Appli
45	209	3.8	9807	10	US-11-097-143-7664	Sequence 7664, Ap

ALIGNMENTS

RESULT 1
US-09-828-068-1
; Sequence 1, Application US/09828068
; Patent No. US20020157137A1
; GENERAL INFORMATION:
; APPLICANT: Moon, Yong-Hwan
; APPLICANT: Chen, Lingjing
; APPLICANT: Sung, Zinmay R.
; TITLE OF INVENTION: NUCLEIC ACIDS THAT CONTROL REPRODUCTIVE DEVELOPMENT IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 018941-001400US
; CURRENT APPLICATION NUMBER: US/09/828,068
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3896
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (241)..(3411)
US-09-828-068-1

Alignment Scores:

Pred. No.:	0	Length:	3896
Score:	5526.00	Matches:	1057
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-09-828-068-2 (1-1057) x US-09-828-068-1 (1-3896)

```
Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
Db        241 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 300

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
Db        301 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 360

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
Db        361 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 420

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
Db        421 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 480

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
Db        481 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 540

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
Db        541 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 600

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
Db        601 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 660

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
Db        661 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 720

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
Db        721 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 780

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
Db        781 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 840

Qy        201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
Db        841 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 900

Qy        221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
Db        901 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 960

Qy        241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
Db        961 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1020

Qy        261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
Db       1021 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1080

Qy        281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
Db       1081 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1140

Qy        301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
Db       1141 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1200

Qy        321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
Db       1201 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1260

Qy        341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
```

```
Db       1261 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1320

Qy        361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
Db       1321 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1380

Qy        381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
Db       1381 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1440

Qy        401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
Db       1441 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1500

Qy        421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db       1501 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1560

Qy        441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
Db       1561 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 1620

Qy        461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db       1621 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 1680

Qy        481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
Db       1681 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 1740

Qy        501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
Db       1741 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 1800

Qy        521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
Db       1801 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 1860

Qy        541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
Db       1861 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 1920

Qy        561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
Db       1921 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 1980

Qy        581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
Db       1981 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2040

Qy        601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db       2041 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2100

Qy        621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db       2101 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2160

Qy        641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db       2161 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2220

Qy        661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db       2221 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2280

Qy        681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db       2281 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2340

Qy        701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
Db       2341 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2400
```

```
Qy        721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2460

Qy        741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2520

Qy        761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2580

Qy        781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 2640

Qy        801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 2700

Qy        821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 2760

Qy        841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 2820

Qy        861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 2880

Qy        881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 2940

Qy        901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3000

Qy        921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3060

Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3120

Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3180

Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3240

Qy       1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3300

Qy       1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3360

Qy       1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3411

RESULT 2
US-10-447-135-1
; Sequence 1, Application US/10447135
; Publication No. US20030199684A1
; GENERAL INFORMATION:
;  APPLICANT: Hirochika, Hirohiko
;  APPLICANT:  Yamazaki, Muneo
;  APPLICANT:  Miyao, Akio
;  TITLE OF INVENTION: A novel gene involved in brassinosteroid responses
;  FILE REFERENCE: MAFF-1 DIV
;  CURRENT APPLICATION NUMBER: US/10/447,135
;  CURRENT FILING DATE:  2003-05-27
;  PRIOR APPLICATION NUMBER: US 09/721,114
;  PRIOR FILING DATE: 2000-11-22
;  PRIOR APPLICATION NUMBER: JAPAN 2000-149106
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4310
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (655)..(3828)
US-10-447-135-1

Alignment Scores:
Pred. No.:                0              Length:          4310
Score:                    5526.00        Matches:         1057
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       6              Gaps:            0

US-09-828-068-2 (1-1057) x US-10-447-135-1 (1-4310)

Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 714

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 774

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 834

Qy         61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 894

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 954

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1014

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1074

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1134

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1194

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1254
```

```
Qy      201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
Db     1255 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1314

Qy      221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
Db     1315 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1374

Qy      241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
Db     1375 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1434

Qy      261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
Db     1435 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1494

Qy      281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
Db     1495 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1554

Qy      301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
Db     1555 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1614

Qy      321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
Db     1615 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 1674

Qy      341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
Db     1675 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 1734

Qy      361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
Db     1735 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 1794

Qy      381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
Db     1795 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 1854

Qy      401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
Db     1855 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 1914

Qy      421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db     1915 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 1974

Qy      441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
Db     1975 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2034

Qy      461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db     2035 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2094

Qy      481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
Db     2095 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2154

Qy      501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
Db     2155 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2214

Qy      521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
Db     2215 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2274

Qy      541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
Db     2275 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2334

Qy      561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
Db     2335 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2394

Qy      581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
Db     2395 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2454

Qy      601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db     2455 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2514

Qy      621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db     2515 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2574

Qy      641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db     2575 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2634

Qy      661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db     2635 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 2694

Qy      681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db     2695 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 2754

Qy      701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
Db     2755 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 2814

Qy      721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
Db     2815 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 2874

Qy      741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
Db     2875 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 2934

Qy      761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
Db     2935 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 2994

Qy      781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
Db     2995 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3054

Qy      801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
Db     3055 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3114

Qy      821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
Db     3115 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3174

Qy      841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
Db     3175 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3234

Qy      861 MetAspProSerThrLeuAlaSerProProAsnTyrGlyThrSerSerArgAsnGlnMet 880
Db     3235 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3294

Qy      881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
Db     3295 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3354

Qy      901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
Db     3355 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3414

Qy      921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
```

```
Db       3415 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3474

Qy        941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3534

Qy        961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3594

Qy        981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3595 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3654

Qy       1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3655 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 3714

Qy       1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3715 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 3774

Qy       1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3775 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 3825

RESULT 3
US-10-437-963-40889
; Sequence 40889, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 40889
;   LENGTH: 4640
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_44288C.1
US-10-437-963-40889

Alignment Scores:
Pred. No.:                0              Length:          4640
Score:                    5526.00        Matches:         1057
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       7              Gaps:            0

US-09-828-068-2 (1-1057) x US-10-437-963-40889 (1-4640)

Qy           1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         985 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 1044

Qy          21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1045 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 1104

Qy          41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1105 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 1164

Qy          61 IleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1165 ATAAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 1224

Qy          81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1225 ATTTTCCATGACCAGAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 1284

Qy         101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1285 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 1344

Qy         121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1345 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 1404

Qy         141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1405 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 1464

Qy         161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1465 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 1524

Qy         181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1525 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 1584

Qy         201 AspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnVal 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1585 GATTCACCAATGAAAGATTTGCAAGGGCCAGCCCAAAATTATGATGTGGCAGCAAATGTC 1644

Qy         221 SerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGlnIleThrTrp 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1645 TCTGAGGACAACACTTCTGTTGATGTTGGGGCTTTACCTGAAGTTCCCCAGATTACATGG 1704

Qy         241 HisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValVal 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1705 CACATAGAAGTAAATGGTGCAGATCAACCTCCATCCACTCCAAAACTTTCTGAAGTGGTC 1764

Qy         261 LeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCys 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1765 CTCAAAAGAAATGAAGATGAAAATGGAAAAACTGAAGAGACTCTTGTTGCTGAGCAGTGC 1824

Qy         281 AsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGln 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1825 AATTTGACCAAAGATCCTAACCCAATGTCTGGAAAGGAACGTGATCAGGTTGCTGAGCAG 1884

Qy         301 CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsn 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1885 TGCAATTTGACCAAAGATCCGAAACCAGTGTCTGGGCAGAAATGTGAGCAGATCTGCAAT 1944

Qy         321 GluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLys 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1945 GAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGCTCCAAATCTAAGAGGAAGACGGATAAG 2004

Qy         341 LysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAsp 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2005 AAGTTGATGAAGAAGCAGCAGCACAGCAAGAAACGCACTGCCCAGGCTGATGTTTCAGAT 2064

Qy         361 AlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAla 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2065 GCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTGCGGCTTCTATCAGAAATTATAAATGCT 2124

Qy         381 AsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspProCys 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2125 AACCAGGTTGAGGATTCTAGAAGTGACGAAGTTCATCGTGAAAATGCCGCTGATCCCTGT 2184

Qy         401 GluAspAspArgSerThrIleProValProMetGluValSerMetAspIleProValSer 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2185 GAGGATGATAGAAGTACCATCCCGGTCCCGATGGAAGTAAGCATGGATATTCCTGTTAGC 2244
```

```
Qy       421 AsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyr 440
Db      2245 AACCATACAGTGGGAGAAGATGGGTTAAAATCAAGTAAGAACAAGACAAAACGCAAATAC 2304

Qy       441 SerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
Db      2305 TCTGATGTTGTAGATGATGGATCATCACTTATGAACTGGCTGAATGGAAAAAAGAAAAGA 2364

Qy       461 ThrGlySerValHisHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysVal 480
Db      2365 ACTGGAAGTGTGCATCACACAGTTGCTCATCCAGCTGGGAATTTGAGCAACAAAAAAGTG 2424

Qy       481 ThrProThrAlaSerThrGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThr 500
Db      2425 ACACCCACTGCGAGTACTCAGCATGATGATGAGAATGATACTGAAAATGGTCTTGACACA 2484

Qy       501 AsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSer 520
Db      2485 AATATGCATAAGACAGATGTCTGTCAGCATGTATCAGAAATCTCCACACAGAGGTGCTCA 2544

Qy       521 SerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLys 540
Db      2545 TCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGGAAAACACATTCAGCTGCTAGTACCAAA 2604

Qy       541 TyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGln 560
Db      2605 TATGGTGGTGAAAGCACCAGAAATGGTCAGAACATACATGTACTCAGCGCAGAAGATCAA 2664

Qy       561 CysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAlaGlu 580
Db      2665 TGCCAGATGGAAACCGAAAACTCTGTTCTGAGTCACTCGGCAAAGGTTTCTCCAGCTGAG 2724

Qy       581 HisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLys 600
Db      2725 CATGATATCCAAATTATGTCTGACCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAG 2784

Qy       601 GlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIle 620
Db      2785 CAAAAACTTGAAGTGACTCGTGAAAAACAGACCATGATAGATGACATCCCCATGGATATT 2844

Qy       621 ValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSer 640
Db      2845 GTTGAACTGCTAGCTAAAAACCAGCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCT 2904

Qy       641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db      2905 GACATCAACCGTATTCAATCCAAGACAACTGCTGATGATGATTGTGTAATAGTAGCTGCC 2964

Qy       661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGlnLysSerLeu 680
Db      2965 AAGGATGGTTCAGATTATGCATCAAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTG 3024

Qy       681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db      3025 GCATCCCAAAGTACACAGAAGGAGTTACAGGGTCATTTGGCATTGACCACACAAGAGTCT 3084

Qy       701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGlu 720
Db      3085 CCACATCCTCAGAACTTTCAGTCTACTCAGGAACAGCAGACACATTTGCGGATGGAAGAA 3144

Qy       721 MetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAla 740
Db      3145 ATGGTCACTATTGCTGCAAGCTCACCACTATTTTCACATCATGATGATCAGTATATTGCT 3204

Qy       741 GluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPhe 760
Db      3205 GAAGCACCAACTGAACATTGGGGCCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTT 3264

Qy       761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIle 780
Db      3265 AAGGCCACTACAAGAAATTCTCCAGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATC 3324

Qy       781 GlnAlaValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArg 800
Db      3325 CAAGCAGTTGACTTGACTTCTACTCATGTCATGGGATCTTCCAGCAATTATGCATCTCGC 3384

Qy       801 GlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAla 820
Db      3385 CAACCAGTAATTGCGCCACTGGACCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCA 3444

Qy       821 ArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsn 840
Db      3445 AGAAATTTTCCAAGCACAATAGCAACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAAT 3504

Qy       841 AlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMet 860
Db      3505 GCTGGACAAGTAGTCTTGTATCCTAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATG 3564

Qy       861 MetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMet 880
Db      3565 ATGGATCCATCAACATTAGCAAGCTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATG 3624

Qy       881 GluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThr 900
Db      3625 GAGTCTCAACTTCATAATTCTCAGTATGCACATAATCAGTACAAAGGATCAACCAGCACA 3684

Qy       901 SerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHis 920
Db      3685 TCATATGGCAGTAACCTGAATGGAAAGATTCCATTGACATTCGAAGACTTATCACGGCAT 3744

Qy       921 GlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySer 940
Db      3745 CAGCTGCATGATCTGCACAGACCTTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCC 3804

Qy       941 LeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLys 960
Db      3805 TTGCTGCAGAAGGAAATTGCAAACTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAG 3864

Qy       961 LeuGlyValSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAla 980
Db      3865 TTAGGAGTGTCAACAGGAATAACATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCC 3924

Qy       981 LeuAsnSerGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSer 1000
Db      3925 CTGAATTCTGGAATGTTTTCAGCAAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCC 3984

Qy      1001 SerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGly 1020
Db      3985 AGTGCAGATTTTTTATCAGCGAGGAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGT 4044

Qy      1021 LysMetValHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsn 1040
Db      4045 AAAATGGTTCATCCCTTGGATCGGTTTGTGAGACAGGATATCTGTATAACTAACAAGAAC 4104

Qy      1041 ProAlaAspPheThrThrIleSerAsnAspAsnGluTyrMetAspTyrArg 1057
Db      4105 CCAGCTGATTTTACTACAATCAGTAACGATAACGAGTATATGGATTACCGC 4155
```

RESULT 4
US-10-447-135-3
; Sequence 3, Application US/10447135
; Publication No. US20030199684A1
; GENERAL INFORMATION:
; APPLICANT: Hirochika, Hirohiko
; APPLICANT: Yamazaki, Muneo
; APPLICANT: Miyao, Akio
; TITLE OF INVENTION: A novel gene involved in brassinosteroid responses
; FILE REFERENCE: MAFF-1 DIV
; CURRENT APPLICATION NUMBER: US/10/447,135
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: US 09/721,114
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JAPAN 2000-149106
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 3

```
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 9455
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-10-447-135-3

Alignment Scores:
Pred. No.:              0              Length:          9455
Score:                  5191.00        Matches:         1055
Percent Similarity:     68.1%          Conservative:    0
Best Local Similarity:  68.1%          Mismatches:      2
Query Match:            93.9%          Indels:          495
DB:                     6              Gaps:            3

US-09-828-068-2 (1-1057) x US-10-447-135-3 (1-9455)

Qy          1 MetGluIleValAlaValAspGlnGluGlyAlaArgValValGlyThrAsnCysMetLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4223 ATGGAGATTGTTGCAGTAGATCAGGAGGGAGCTCGTGTTGTTGGGACGAACTGTATGCTT 4282

Qy         21 AlaArgGlyGlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGln 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4283 GCTCGTGGTGGAACTGGTGCTGTAGCGCCAGTGTTGGAGCTGACAGCGACGCCTCGTCAG 4342

Qy         41 AspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPheSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4343 GATGCAGCCGCTGAAGCTGGTGTAGACGAACCGGCACAACACCAATGCGAGCATTTCTCC 4402

Qy         61 Ile---------------------------------------------------------- 61
Db       4403 AT-AAGGTAATCATTTTCTGTATTTCCAATTCCAGTATCGCGTTGTGGATGAATAATGAA 4461

Qy         61 ---------------------------------------------------------- 61
Db       4462 TCGGCATGTCATGCCATATTGCACTGTTTGATGGAAGAGTATGATTGATCGTGGTTTTTG 4521

Qy         61 ---------------------------------------------------------- 61
Db       4522 CACAGTTTGCTGTTGGGACTTATATGGTCATCTGTTTTGTACGATCGTATACACTGGGTC 4581

Qy         61 ---------------------------------------------------------- 61
Db       4582 GACATGCTTATGACTTTGGTTCGATTTAGGAAGTCAATACATCCACTACTAGCTCTATAT 4641

Qy         61 ---------------------------------------------------------- 61
Db       4642 CTAGCCATGTGAACTCATTTATGCCATAGCACAGCTAGCAGGCTAGCAGCAAAAAATATA 4701

Qy         61 ---------------------------------------------------------- 61
Db       4702 TATAATATTTGCATATATGTTGGTGTTTCATGTATCTTTATACTCTACGTACATCCATTA 4761

Qy         61 ---------------------------------------------------------- 61
Db       4762 ATATCTTCAATGTATGAATCTGAGCACATGATTGTGAGTGCTACACATATGCATGTCTGT 4821

Qy         61 ---------------------------------------------------------- 61
Db       4822 ATGTGTGTTCATTAGGTGTTTGATCATATTTGTTTGTGTTGGGGTGCGCATGCATTTATT 4881

Qy         61 ---------------------------------------------------------- 61
Db       4882 CAGGCCATGCTGTAGGCTGTAGCTAGATATTTGTGTTTGTATATTATTTCTGTTGAACAA 4941

Qy         61 ---------------------------------------------------------- 61
Db       4942 GCTGATTACTAATGAAATTAACCTTTTTGGGGTACACTCATATATTGGGCCCTACATTTT 5001

Qy         61 ---------------------------------------------------------- 61
Db       5002 TGTAATCATTTTTCCTTTGTGCTGAGGTTCAGCATAAAACTTTTTTATCATAAGCATGTT 5061

Qy         61 ---------------------------------------------------------- 61
Db       5062 TACATCCTAGGAGATTCTTAGAACTGATGGTTTCTTCATATTTGCATTATGTTTGATTTG 5121

Qy         61 ---------------------------------------------------------- 61
Db       5122 ATAGTCCATTATTATTTTAAGCCTTTTCAATTGTTTAGAGATTCTAGAGATGATATATAT 5181

Qy         61 ---------------------------------------------------------- 61
Db       5182 CAACCATAGACTTGTCACGTTTTGGTTTAATACTTTCTAGAACTAATTAGATTATTATTT 5241

Qy         61 ---------------------------------------------------------- 61
Db       5242 TTGTAGTTTATCCTGTCATGCTATTTGTATTATCTTTGAATTCAAACTGCAATACTTAGA 5301

Qy         61 ---------------------------------------------------------- 61
Db       5302 TTATCTTGAAGGTCCTCTTTTTCTGGACTGTACAAGCTATGTATGAAATGCCTACCTCCC 5361

Qy         61 ---------------------------------------------------------- 61
Db       5362 AGCATCCTTTAGATTATGTAGGGCCTTTTCTGAGTTTATCAGTTGTATATTGACTGAAGC 5421

Qy         61 ---------------------------------------------------------- 61
Db       5422 ACGCAATGTGCTATATATATGTGCCATGCATCTTTATAATGATAATCTTATTTTTCTTGT 5481

Qy         62 ---ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5482 ACCAGAGGGTATGTTGCTCTTCTTCAGAAGAAGGATCCAAAATTCTGCTCTCTATCTCGG 5541

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5542 ATTTTCCATGACCAGAAAAAAATGTGATGAACACAAAGCTAGTTCAAGCCCATTTTCTGTA 5601

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5602 GCAAAGTTTCGACGATGGGATTGCTCGAAGTGCTTGGATAAGTTGAAAACTTCAGATAAT 5661

Qy        121 GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5662 GGAACAGCACCAAGAACTCTTCCCGCAAAGCAGAATGGCACAAGTGATGGTTGCTCCATC 5721

Qy        141 ThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSerProSer 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5722 ACATTTGTTCGGAGCACTTTTGTGCCTGCTAGTGTTGGTTCCCAAAAAGTGTCTCCTAGC 5781

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5782 ACACAATCATCTCAAGGGAAGAATGCTGATAGATCAACTCTTCCAAAGAGTGTGCAAGAA 5841

Qy        181 GlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5842 GGCAATGACTCCAAATGCAATGCGCCTTCTGGCAAGAATGGAGCTGCTGAGGCCAATACT 5901

Qy        201 AspSerProMetLys---------------------------------------------- 205
              ||||||||||||
Db       5902 GATTCACCAATGAA-AGGTATGGTAGATGTAGAGCCTTTCAAATTCCTAAGTAGGATTTT 5960

Qy        206 ----------------------------------------AspLeuGlnGlyProAla 211
                                                     ||||||||||||||||||
Db       5961 ATTTAAGGTATAGAATAAACTAATGTTTGTGTGATTTTCTCAGATTTGCAAGGGCCAGCC 6020

Qy        212 GlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAla 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6021 CAAAATTATGATGTGGCAGCAAATGTCTCTGAGGACAACACTTCTGTTGATGTTGGGGCT 6080

Qy        232 LeuProGluValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6081 TTACCTGAAGTTCCCCAGATTACATGGCACATAGAAGTAAATGGTGCAGATCAACCTCCA 6140

Qy        252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
```

```
Db      6141 TCCACTCCAAAACTTTCTGAAGTGGTCCTCAAAAGAAATGAAGATGAAAATGGAAAAACT 6200
Qy       272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
Db      6201 GAAGAGACTCTTGTTGCTGAGCAGTGCAATTTGACCAAAGATCCTAACCCAATGTCTGGA 6260
Qy       292 LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
Db      6261 AAGGAACGTGATCAGGTTGCTGAGCAGTGCAATTTGACCAAAGATCCGAAACCAGTGTCT 6320
Qy       312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
Db      6321 GGGCAGAAATGTGAGCAGATCTGCAATGAGCCATGTGAAGAAGTTGTTCTCAAAAGAAGC 6380
Qy       332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
Db      6381 TCCAAATCTAAGAGGAAGACGGATAAGAAGTTGATGAAGAAGCAGCAGCACAGCAAGAAA 6440
Qy       352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
Db      6441 CGCACTGCCCAGGCTGATGTTTCAGATGCAAAGCTTTGTCGGAGAAAGCCAAAAAAGGTG 6500
Qy       372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVal 391
Db      6501 CGGCTTCTATCAGAAATTATAAATGCTAACCAGGTTGAGGATTCTAGAAGTGACGAAGTT 6560
Qy       392 HisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProValProMet 411
Db      6561 CATCGTGAAAATGCCGCTGATCCCTGTGAGGATGATAGAAGTACCATCCCGGTCCCGATG 6620
Qy       412 GluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSer 431
Db      6621 GAAGTAAGCATGGATATTCCTGTTAGCAACCATACAGTGGGAGAAGATGGGTTAAAATCA 6680
Qy       432 SerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMet 451
Db      6681 AGTAAGAACAAGACAAAACGCAAATACTCTGATGTTGTAGATGATGGATCATCACTTATG 6740
Qy       452 AsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHisPro 471
Db      6741 AACTGGCTGAATGGAAAAAAGAAAAGAACTGGAAGTGTGCATCACACAGTTGCTCATCCA 6800
Qy       472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAspGlu 491
Db      6801 GCTGGGAATTTGAGCAACAAAAAAGTGACACCCACTGCGAGTACTCAGCATGATGATGAG 6860
Qy       492 AsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVal 511
Db      6861 AATGATACTGAAAATGGTCTTGACACAAATATGCATAAGACAGATGTCTGTCAGCATGTA 6920
Qy       512 SerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGly 531
Db      6921 TCAGAAATCTCCACACAGAGGTGCTCATCAAAGGGGAAAACAGCGGGTTTGAGTAAGGGG 6980
Qy       532 LysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsn 551
Db      6981 AAAACACATTCAGCTGCTAGTACCAAATATGGTGGTGAAAGCACCAGAAATGGTCAGAAC 7040
Qy       552 IleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSer 571
Db      7041 ATACATGTACTCAGCGCAGAAGATCAATGCCAGATGGAAACCGAAAACTCTGTTCTGAGT 7100
Qy       572 HisSerAlaLys---------------------------------------------- 575
Db      7101 CACTCGGCAAAGGTACGAATTTTGTGAATCATGAGGAATTTTTGCTTTTTAAATTGACTG 7160
Qy       575 ---------------------------------------------------------- 575
Db      7161 AATCAACATTTATCTGTATGAAGGAATAATATTGGTGCATAACAATGTTAAGAAATATGC 7220
Qy       575 ---------------------------------------------------------- 575
```

```
Db      7221 ATACAATGTTTATTTATATGCTTTCCACTGTTCTTCTTTACTTATGTTTTGATACTCTTT 7280
Qy       575 ---------------------------------------------------------- 575
Db      7281 TTGTGTGTGCGTGCATGTGTGCATGTGTGTGTGTGTGTGTGTGCGCGCGTGTGTGTGC 7340
Qy       575 ---------------------------------------------------------- 575
Db      7341 ACGTGCGTGGCGCAATATTCTTTTTTAGACTCATATTATAGTGATTGTAATGGACTGACA 7400
Qy       576 ----------------------ValSerProAlaGluHisAspIleGlnIleMetSerAs 588
Db      7401 TTTTCCTCATTTCTCATCTCAGGTTTCTCCAGCTGAGCATGATATCCAAATTATGTCTGA 7460
Qy       588 pLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGl 608
Db      7461 CCTTCATGAGCAGAGTCTACCCAAGAAGAAAAAGAAGCAAAAACTTGAAGTGACTCGTGA 7520
Qy       608 uLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAlaLysAsnGl 628
Db      7521 AAAACAGACCATGATAGATGACATCCCCATGGATATTGTTGAACTGCTAGCTAAAAACCA 7580
Qy       628 nHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLy 648
Db      7581 GCATGAGAGGCAGCTTATGACTGAGACTGATTGTTCTGACATCAACCGTATTCAATCCAA 7640
Qy       648 sThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe 668
Db      7641 GACAACTGCTGATGATGATTGTGTAATAGTAGCTGCCAAGGATGGTTCAGATTATGCATC 7700
Qy       668 rSerValPheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGl 688
Db      7701 AAGTGTGTTTGACACTAATTCCCAACAGAAGTCCTTGGCATCCCAAAGTACACAGAAGGA 7760
Qy       688 uLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSe 708
Db      7761 GTTACAGGGTCATTTGGCATTGACCACACAAGAGTCTCCACATCCTCAGAACTTTCAGTC 7820
Qy       708 rThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSe 728
Db      7821 TACTCAGGAACAGCAGACACATTTGCGGATGGAAGAAATGGTCACTATTGCTGCAAGCTC 7880
Qy       728 rProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrpGl 748
Db      7881 ACCACTATTTTCACATCATGATGATCAGTATATTGCTGAAGCACCAACTGAACATTGGGG 7940
Qy       748 yArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSerPr 768
Db      7941 CCGTAAGGACGCAAAGAAGCTAACGTGGGAGCAATTTAAGGCCACTACAAGAAATTCTCC 8000
Qy       768 oAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSerTh 788
Db      8001 AGCAGCAACATGTGGTGCTCAATTTAGACCTGGTATCCAAGCAGTTGACTTGACTTCTAC 8060
Qy       788 rHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAs 808
Db      8061 TCATGTCATGGGATCTTCCAGCAATTATGCATCTCGCCAACCAGTAATTGCGCCACTGGA 8120
Qy       808 pArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAl 828
Db      8121 CCGCTATGCTGAAAGAGCGGTTAACCAGGTCCATGCAAGAAATTTTCCAAGCACAATAGC 8180
Qy       828 aThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrPr 848
Db      8181 AACCATGGAAGCGAGTAAGTTATGTGATCGGAGAAATGCTGGACAAGTAGTCTTGTATCC 8240
Qy       848 oLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSe 868
Db      8241 TAAAGAATCCATGCCTGCGACGCATCTTCTGAGAATGATGGATCCATCAACATTAGCAAG 8300
Qy       868 rPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGl 888
Db      8301 CTTCCCCAACTATGGAACTTCTAGCAGGAACCAGATGGAGTCTCAACTTCATAATTCTCA 8360
```

```
Qy       888 nTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGl 908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8361 GTATGCACATAATCAGTACAAAGGATCAACCAGCACATCATATGGCAGTAACCTGAATGG 8420

Qy       908 yLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgPr 928
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8421 AAAGATTCCATTGACATTCGAAGACTTATCACGGCATCAGCTGCATGATCTGCACAGACC 8480

Qy       928 oLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIleAlaAs 948
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8481 TTTACGCCCACATCCTAGAGTTGGTGTGCTTGGCTCCTTGCTGCAGAAGGAAATTGCAAA 8540

Qy       948 nTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleTh 968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8541 CTGGTCGGAGAACTGTGGCACACAATCTGGTTATAAGTTAGGAGTGTCAACAGGAATAAC 8600

Qy       968 rSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAl 988
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8601 ATCGCATCAGATGAACAGAAAGGAACATTTTGAAGCCCTGAATTCTGGAATGTTTTCAGC 8660

Qy       988 aLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAspPheLeuSerAlaAr 1008
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8661 AAAATGGAATGCATTGCAGTTGGGTTCTGTTAGCTCCAGTGCAGATTTTTTATCAGCGAG 8720

Qy      1008 gAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHisProLeuAspAr 1028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8721 GAACAGCATAGCTCAATCTTGGACCAGAGGCAAGGGTAAAATGGTTCATCCCTTGGATCG 8780

Qy      1028 gPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSe 1048
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8781 GTTTGTGAGACAGGATATCTGTATAACTAACAAGAACCCAGCTGATTTTACTACAATCAG 8840

Qy      1048 rAsnAspAsnGluTyrMetAspTyrArg 1057
             |||||||||||||||||||||||||||
Db      8841 TAACGATAACGAGTATATGGATTACCGC 8868

RESULT 5
US-10-425-114-4042
; Sequence 4042, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 4042
;   LENGTH: 871
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700347436_FLI
US-10-425-114-4042

Alignment Scores:
Pred. No.:              5.07e-33        Length:         871
Score:                  491.50          Matches:        107
Percent Similarity:     65.8%           Conservative:   18
Best Local Similarity:  56.3%           Mismatches:     50
Query Match:            8.9%            Indels:         15
DB:                     7               Gaps:           6

US-09-828-068-2 (1-1057) x US-10-425-114-4042 (1-871)

Qy       871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu------SerGlnLeuHisAsnSerGln 888
```

```
             ||||||      :::::|||   |||||||||      :::|||         ||||||
Db         2 AACTACCAAAGAGCTAATAGGCGGCAGATGGAGCTTCAAACGCAAGCTCTCGGCTCGCAG 61

Qy       889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907
             |||      |||      |||      |||      ||||||||||||||||||
Db        62 TACACACAGCATGAGCATTACAATGCGTCACCGAGCACGTCATATGGAAGCCAA------ 115

Qy       908 GlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926
             |||||||||   ||||||||||||   ||||||         ||||||
Db       116 ---------CCGCTGACACTGGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGCGC 166

Qy       927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db       167 AGGCCTTTACGCCCTCACCCTCGTGTCGGTGTGCTCGGTTCATTGCTGCAGCAGGAGATC 226

Qy       947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGly 966
             ||||||||||||   |||||||||   ||||||||||||:::||||||         |||
Db       227 GCAAACTGGTCTGGGAACTGCGGGCCGCAGTCTGGGTACAGACTAGGTGATTGTAAAGGG 286

Qy       967 IleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPhe 986
                         ||||||      ::::::|||   |||:::
Db       287 ACAAGGCCGCCTGATGCCAACAGAGCAGGAAACTACGAGACCTTGAGC------------ 334

Qy       987 SerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla------AspPhe 1004
             ||||||   ||||||   ||||||||||||||||||||||:::|||      ::::::
Db       335 TCGGCAGGATGGAACACCCTGCAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGTAC 394

Qy      1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024
                              |||   |||   |||:::   :::|||
Db       395 CGGTTTCCATGGTATGGTACGGGTCAGCCTTCGACCAGTGGCAATGGGAGGACCATTCAT 454

Qy      1025 ProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPhe 1044
             |||||||||:::   ||||||:::||||||||:::||||||:::|||||||||||||||||
Db       455 CCGTTGGATAAGCTCGTGAGGAAGGATATCTGTGTGACTAACAGAAACCCAGCCGATTTC 514

Qy      1045 ThrThrIleSerAsnAspAsnGluTyrMet 1054
             |||   ||||||:::   ||||||||||||
Db       515 ACTGTAATTAGTGACAAGAACGAGTACATG 544

RESULT 6
US-10-425-115-140777
; Sequence 140777, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 140777
;   LENGTH: 884
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_59872C.1
US-10-425-115-140777

Alignment Scores:
Pred. No.:              5.17e-33        Length:         884
Score:                  491.50          Matches:        107
Percent Similarity:     65.8%           Conservative:   18
Best Local Similarity:  56.3%           Mismatches:     50
Query Match:            8.9%            Indels:         15
DB:                     8               Gaps:           6

US-09-828-068-2 (1-1057) x US-10-425-115-140777 (1-884)

Qy       871 AsnTyrGlyThrSerSerArgAsnGlnMetGlu------SerGlnLeuHisAsnSerGln 888
```

```
Db         14 AACTACCAAAGAGCTAATAGGCGGCAGATGGAGCTTCAAACGCAAGCTCTCGGCTCGCAG 73

Qy        889 TyrAla---HisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsn 907

Db         74 TACACACAGCATGAGCATTACAATGCGTCACCGAGCACGTCATATGGAAGCCAA------ 127

Qy        908 GlyLysIleProLeuThrPheGluAspLeuSer---ArgHisGlnLeuHisAspLeuHis 926

Db        128 ---------CCGCTGACACTGGAAGACTTGTCTCGGCGTCATTTCCAGCAAGACTTGCGC 178

Qy        927 ArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGlnLysGluIle 946

Db        179 AGGCCTTTACGCCCTCACCCTCGTGTCGGTGTGCTCGGTTCATTGCTGCAGCAGGAGATC 238

Qy        947 AlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGly 966

Db        239 GCAAACTGGTCTGGGAACTGCGGGCCGCAGTCTGGGTACAGACTAGGTGATTGTAAAGGG 298

Qy        967 IleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPhe 986

Db        299 ACAAGGCCGCCTGATGCCAACAGAGCAGGAAACTACGAGACCTTGAGC------------ 346

Qy        987 SerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAla------AspPhe 1004

Db        347 TCGGCAGGATGGAACACCCTGCAGTTGGGTTCTGTTAGCTCTGCTGCCAATCTGGAGTAC 406

Qy       1005 LeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetValHis 1024

Db        407 CGGTTTCCATGGTATGGTACGGGTCAGCCTTCGACCAGTGGCAATGGGAGGACCATTCAT 466

Qy       1025 ProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPhe 1044

Db        467 CCGTTGGATAAGCTCGTGAGGAAGGATATCTGTGTGACTAACAGAAACCCAGCCGATTTC 526

Qy       1045 ThrThrIleSerAsnAspAsnGluTyrMet 1054

Db        527 ACTGTAATTAGTGACAAGAACGAGTACATG 556

RESULT 7
US-10-425-115-176502
; Sequence 176502, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 176502
;   LENGTH: 1039
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_92562C.1
US-10-425-115-176502

Alignment Scores:
Pred. No.:              1.79e-31        Length:         1039
Score:                  475.00          Matches:        141
Percent Similarity:     49.5%           Conservative:   56
Best Local Similarity:  35.4%           Mismatches:     129
Query Match:            8.6%            Indels:         72
DB:                     8               Gaps:           17

US-09-828-068-2 (1-1057) x US-10-425-115-176502 (1-1039)

Qy         81 IlePheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal 100

Db          4 ATTTTCCATAACCAGCCACAATATGATGAGCACCAT---GATTCGTCCCCAGTTTTGGTG 60

Qy        101 AlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAspAsn 120

Db         61 TCAAAGTATCATCGATGGGATTGCTCAAGATGTCTTGATAGTGTGGAAGTTTCAGGCCAT 120

Qy        121 -GlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIl 140

Db        121 AGGCCAACACTCGGAAAATGTTTCTATGCAGCAGAATGGAATGAATGATGGCTGCTCTAT 180

Qy        140 eThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysVal---SerPr 159

Db        181 TTCAATTGTTCGG------ATTTTGCCTAATAGTGTTGATTCCAGAAGACTGTTTTCTTG 234

Qy        159 oSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGl 179

Db        235 CACACAGCAATCATCTCAAGGGAAT------GATCGATTGACCCTTTCAAAGACTGCCCA 288

Qy        179 nGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn---------------- 193

Db        289 AGAATGCAAT---TCAAAATGCAGTTCTCCTGGCAACAAAGCAATTACTGCGATGAATGT 345

Qy        194 -GlyAlaAlaGluAlaAsnThrAspSerProMetLysAsp------------------- 206

Db        346 TCCAGTAGCTGAAGAAAATGTTCCAGAGGCACTCGTTGATACGAGGGTTCCATCTATAGA 405

Qy        207 ----LeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnTh 225

Db        406 AGTTTTACAGGCCTCTCCCAATAGTATTGATTTGTCAGCAAACATCTTG---AATGCTGT 462

Qy        225 rSerValAspValGlyAlaLeuProGluValProGlnIleThrTrpHisIleGluValAs 245

Db        463 CTCGAAGAATGTCCGCGATTTACCTGAT-----------------GATGTTGAAGAAAA 504

Qy        245 nGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeuLysArgAsnGl 265

Db        505 TGGCACACAGAATCCACACAGTCTA-----------AAACCTTGTGTGGCGCCAAATGA 552

Qy        265 uAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAs 285

Db        553 AGATGAGAATAACATAGCTAATGATGACCTC----------------------AATGA 588

Qy        285 pProAsnProMetSerGlyLysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLy 305

Db        589 GCCTAAT-----------------------------------GTTTGT---------- 601

Qy        305 sAspProLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGl 325

Db        602 -------AAACCAGTATCTGGGCACAAAAGCAAACAAGTCTGCAAC------------- 640

Qy        325 uValValLeuLysArgSerSerLysSerLysArgLysThrAspLysLysLeuMetLysLy 345

Db        641 ----ATGGGTCCACGCCGAGCATCGTCAAAAAGAAATGTTGAATCTGATGGTAAGAGGAA 696

Qy        345 sGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysAr 365

Db        697 AAGG---------AATAAATCTACTAAACTGCCTGGTATTTCAGATCTCAAATTTTGTCA 747

Qy        365 gArgLysPro-LysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluA 385

Db        748 GAGAAAGCCAAAAAAAGACAAGGCTGTTATCAGAACTCATAGGTTCTGAACAGGTAGGGG 807

Qy        385 spSerArg---SerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspA 404

Db        808 TTTCTGCAGATGCCGTTCAAGTGGACCATACTAACAGTGTTGATCTCTGTGAGGGTGGTA 867

Qy        404 rgSerThrIleProValProMetGluValSerMetAspIleProValSerAsnHisThrV 424

Db        868 AAAGA------AAAATGCACCTTGAGGTTGGAAAAGGTAATGACACTACTAATCAGCAAG 921

Qy        424 alGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValV 444
```

```
Db        922 TGGACGAAATCCAGTCAAGAGCTGTCAAGAACAAGGCAAAATACACAGGAGTAGACAAGT 981

Qy        444 alAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg 460
              :::::|||||||||||||||||||||||||          |||:::
Db        982 TAGAAGATGGATCCTCTCTAATGAACTGGCTGAAAAGCACTAATAAGAAA 1031

RESULT 8
US-10-739-930-66
; Sequence 66, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-21(53377)B
;  CURRENT APPLICATION NUMBER: US/10/739,930
;  CURRENT FILING DATE:  2003-12-18
;  NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 66
;   LENGTH: 3679
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER108532_1
US-10-739-930-66

Alignment Scores:
Pred. No.:              6.59e-17       Length:         3679
Score:                  317.00         Matches:        233
Percent Similarity:     35.8%          Conservative:   176
Best Local Similarity:  20.4%          Mismatches:     391
Query Match:            5.7%           Indels:         342
DB:                     8              Gaps:           54

US-09-828-068-2 (1-1057) x US-10-739-930-66 (1-3679)

Qy         83 HisAspGlnLysLysCysAspGluHis---------LysAlaSerSerSerProPheSer 99
              |||      |||       |||         ||||||       :::
Db        647 CATGGGCCAAAAGATTGTGGACTGCATTCAAACTCAAAAGCTATTGGAAACTCTTCTGTT 706

Qy        100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
              :::          :::::::::             :::|||    |||
Db        707 ATCGAAAGTAAAAGCAAGTTCAATTCGCTAACTATCATTGATCACGAGAAAGAAAAGAAA 766

Qy        120 AsnGlyThrAlaProArgThrLeuProAlaLys------GlnAsnGlyThrSerAspGly 137
                       |||      :::   |||         |||      :::|||
Db        767 ACTGATATTGCAGATAATGCTATTGAGGAGAAAGTGGGTGTAAACTGTGAGAATGATGAT 826

Qy        138 CysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysVal 157
              :::   |||            ||||||:::      :::   |||
Db        827 CAGACAGCTACT------------ACGTTTCTCAAGAAAGCACGTGGTCGACCTATGGGT 874

Qy        158 SerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSer 177
              :::   :::   :::|||      |||                   |||
Db        875 GCTTCTAATGTTAGGAGCAAGAGCAGAAAG------------------CTTGTGAGTCCG 916

Qy        178 ValGlnGluGlyAsnAsp---------SerLysCysAsnAlaProSerGlyLysAsnGly 194
                 |||   ||||||:::         |||   |||   ||||||
Db        917 GAGCAGGTAGGAAACAACAGATCTAAAGAAAAACTAAACAAACCATCG------------ 964

Qy        195 AlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyr 214
                                             |||:::            :::
Db        965 ------------------------------ATGGATATTAGCAGCTGGAAAGAGAAACAA 994

Qy        215 AspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGlu 234
              :::|||         |||:::      :::|||      |||      :::
Db        995 AATGTGGATCAGGCTGTGACAACGTTCGGCTCATCTGAAATTGCTGGTGTGGTT------ 1048

Qy        235 ValProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrPro 254
                                                   |||   ||||||
Db       1049 ------------------------------------GAGGATACACCACCTAAGGCAACC 1072

Qy        255 LysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThr 274
              |||      :::   :::                     |||   :::||||||   :::|||
Db       1073 AAGAATCATAAAGGCATTCGCGGTCTGATGGAATGCGATAACGGGTCATCAGAAAGTATA 1132

Qy        275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
                 :::|||                                 |||||||||   :::|||
Db       1133 AATCTTGCT---------------------------------ATGAGTGGGTTGCAGCGT 1159

Qy        295 ---------AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
                       :::|||            |||      :::      |||
Db       1160 AGGAAATCTCGCAAGGTTCGTCTACTCAGTGAGTTGCTTGGTAATACAAAAACCAGTGGT 1219

Qy        312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
              |||         :::                  ||||||      ||||||:::
Db       1220 GGTAGTAACATCAGAAAA-----------------GAAGAGTCTGCTTTGAAGAAGGAA 1261

Qy        332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHis---SerLys 350
              |||         ||||||      :::||||||:::   :::      :::   |||:::
Db       1262 TCAGTTAGAGGTCGAAAA------AGAAAGTTGTTACCTGAAAACAATTATGTCAGCCGG 1315

Qy        351 LysArgThrAlaGlnAlaAspValSerAsp------AlaLysLeuCys------------ 364
                 :::                  |||:::      :::|||   |||
Db       1316 ATATTGAGTACAATGGGTGCAACCTCTGAAAATGCTTCCAAAAGTTGTGACTCTGATCAA 1375

Qy        365 ---------------------------------------------------ArgArgLys 367
                                                                 :::::::::
Db       1376 GGTAATAGTGAATCAACTGATAGTGGGTTTGACAGAACTCCATTTAAGGGTAAGCAGAGA 1435

Qy        368 ProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg 387
                 ::::::   :::::::::   |||   :::   :::         |||   |||:::
Db       1436 AACAGAAGATTTCAGGTTGTTGACGAGTTTGTACCATCACTTCCTTGTGAAACTTCACAA 1495

Qy        388 SerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThr--- 406
                    |||      :::|||:::   |||||||||   :::      |||||||||
Db       1496 ------GAAGGTATCAAGGAGCATGATGCAGATCCTAGTAAG------AGATCAACTCCT 1543

Qy        407 ------------------------------IleProValProMetGluValSerMetAsp 416
                                            :::|||   |||
Db       1544 GCGCACTCTTTATTCACTGGAAACGATTCTGTTCCTTGTCCT------------------ 1585

Qy        417 IleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThr 436
                 |||   :::         |||      |||      |||         |||   ||||||
Db       1586 ---CCGGGTACTCAGAGAACA------GAGAGGAAGCTCAGTTTACCCAAGAAGAAGACA 1636

Qy        437 LysArgLysTyrSerAspValValAspAspGlySerSerLeuMet------------Asn 452
              |||:::            |||:::|||:::|||   |||   :::            |||
Db       1637 AAGAAG---------CCTGTAATCGATAATGGAAGAGCACTGTGATCAGTTTTAGTAAC 1687

Qy        453 TrpLeuAsnGlyLysLys------LysArgThrGlySerValHisHisThrValAlaHis 470
                 :::::|||   :::            ||||||         :::||||||:::
Db       1688 GGCATTGATGGAAGTCAAGTTAACTCGCATACTGGTCCTTCCATGAACACAGTATCCCAA 1747

Qy        471 ProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspAsp 490
                          |||:::   |||:::|||
Db       1748 ACTCGAGACTTATTGAATGGGAAAAGGGTG------------------------------ 1777

Qy        491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAsp------ValCys 508
                                |||||   |||         :::|||
Db       1778 ---------------GGCGGTTTATTTGACAACCGTTTGGCTTCAGATGGATATTTCAGA 1822

Qy        509 GlnHisValSerGluIleSerThrGlnArgCysSerSer--------------------- 521
              :::::::::|||:::::::::   :::      :::|||
Db       1823 AAATATCTCTCTCAGGTTAATGATAAGCCGATAACATCTTTGCATTTGCAAGACAATGAT 1882

Qy        522 ------LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThr 539
                    :::   :::   |||      :::      :::      |||:::::::|||
Db       1883 TATGTGAGGTCAAGAGACGCGGAACCAAACTGTCTTCGAGATTTAGTTCCTCTTCTAAA 1942

Qy        540 LysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAsp 559
                    |||         |||   |||      :::   ::::::         :::
Db       1943 TCCAGCTCAGGTGGATGGTTGAGAACTGGA------GTAGATATTGTTGACTTCAGAAAC 1996

Qy        560 GlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLysValSerProAla 579
                       :::|||         :::            :::         |||:::
```

```
Db       1997 AACAACCACAATACAAACAGATCGTCTTTCTCGAACTTGAAGCTAAGATACCCCCCTTCT 2056
Qy        580 GluHisAspIleGlnIleMetSerAspLeu---------------------------- 589
Db       2057 TCTACTGAAGTTGCGGATTTATCTCGGGTGCTGCAAAAGGATGCTTCTGGTGCAGATAGA 2116
Qy        590 ----------------------HisGluGlnSerLeuProLysLysLysLysLysGln 601
Db       2117 AAGGGGAAGACTGTTATGGTCCAAGAACATCATGGAGCACCAAGAAGCCAAAGTCACGAT 2176
Qy        602 LysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleVal 621
Db       2177 AGAAAGGAGACTACGACTGAAGAGCAAAACAAC---GATGATATTCCAATGGAGATAGTG 2233
Qy        622 GluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMet---ThrGluThrAspCysSer 640
Db       2234 GAGCTCATGGCCAAAAACCAGTACGAGAGGTGTCTTCCCGACAAAGAAGAAGATGTTAGC 2293
Qy        641 AspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAla 660
Db       2294 AACAAACAGCCATCACAAGAAACAGCACACAAATCCAAGAATGCTCTACTGATTGATCTC 2353
Qy        661 LysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------Lys 678
Db       2354 AATGAAACCTACGATAACGGGATCTCACTTGAGGACAACAACACATCAAGACCACCAAAA 2413
Qy        679 SerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGln 698
Db       2414 CCGTGTAGTAGCAACGCAAGGAGGGAAGAA---------CATTTTCCTATGGGAAGACAG 2464
Qy        699 GluSerProHis------------------------------ProGlnAsn--------- 705
Db       2465 CAGAACTCTCATGACTTCTTCCCAATAAGTCAGCCTTATGTGCCTTCTCCGTTTGGGATC 2524
Qy        706 PheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAla 725
Db       2525 TTTCCTCCTACCCAAGAAAACCGA---------------------------------- 2548
Qy        726 AlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGlu 745
Db       2549 GCCAGCTCCATCCGGTTTTCTGGTCACAACTGTCAGTGGCTTGGGAATTTGCCAACTGTG 2608
Qy        746 HisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db       2609 ------GGTAATCAGAACCCTTCTCCATCCTCATTTCGGGTATTACGTGCT--------- 2653
Qy        766 AsnSerProAlaAlaThrCysGlyAla---------GlnPheArgProGlyIleGlnAla 782
Db       2654 ---------TGTGATACTTGCCAGAGTGTTCCTAATCAATACAGA-------------- 2689
Qy        783 ValAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnPro 802
Db       2690 ---------GAAGCTTCTCATCCAATTTGGCCATCTTCC-------------------- 2719
Qy        803 ValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsn 822
Db       2720 ATGATACCACCACAGAGTCAATACAAGCCAGTTTCTTTAAATATTAATCAGTCAACAAAT 2779
Qy        823 PheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGly 842
Db       2780 ---CCGGGTACGCTTTCA------CAGGCATCTAACAATGAAAATACATGGAACCTCAAC 2830
Qy        843 GlnValValLeuTyrProLysGlu---------------------------------- 850
Db       2831 TTTGTTGCTGCCAACGGGAAGCAAAAATGTGGGCCTAATCCAGAATTTTCATTTGGCTGC 2890
Qy        850 ------------------------------------------------------------ 850
Db       2891 AAACATGCTGCTGGGGTTAGTAGTAGTAGTAGTAGGCCAATAGATAACTTTTCTAGTGAG 2950
Qy        851 ---SerMetProAlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPhe 869
Db       2951 AGCTCTATACCGGCATTGCATCTACTCAGCCTTCTGGATCCTCGCCTGAGGTCAACGACT 3010

Qy        870 Pro------AsnTyrGlyThrSer------SerArgAsnGlnMetGluSerGlnLeuHis 885
Db       3011 CCCGCTGACCAACACGGAAACACTAAATTTACTAAAAGACATTTTCCGCCAGCCAACCAG 3070
Qy        886 AsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsn 905
Db       3071 TCCAAAGAGTTTATAGAGCTTCAAACAGGGGACTCTAGTAAGTCAGCCTACTCAACTAAG 3130
Qy        906 LeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeu 925
Db       3131 ---------CAGATACCT------TTTGATCTATACAGCAAGAGATTCACACAAGAGCCT 3175
Qy        926 HisArgProLeuArgPro---HisProArgValGlyValLeuGlySerLeuLeuGlnLys 944
Db       3176 TCCCGGAAGAGTTTCCCCATCACTCCACCTATTGGGACG------TCTTCACTTTCATTT 3229
Qy        945 GluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSer 964
Db       3230 CAAAATGCTTCATGGAGTCCTCAT---------------------------------- 3253
Qy        965 ThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGly 984
Db       3254 --------------CATCAAGAGAAGAAAACCAAGAGAAAAGACACC------------ 3286
Qy        985 MetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp--- 1003
Db       3287 ---TTTGCTCCTGTTTACAATACTCATGAAAAGCCGGTGTTTGCAAGCAGCAACGACCAA 3343
Qy       1004 ------------PheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLys 1019
Db       3344 GCGAAGTTCCAGCTGCTGGGAGCATCGAATTCC-------------------------- 3376
Qy       1020 GlyLysMetValHisProLeu--------------AspArgPheValArgGlnAsp--- 1033
Db       3377 ------ATGATGCTTCCTTTGAAATTTCACATGACGGATAAAGAAAAGAAACAAAAGAGA 3430
Qy       1033 ------------------------------------------------------------ 1033
Db       3431 AAAGCAGAGAGCTGCAATAACAATGCCTCTGCGGGACCTGTGAAGAACAGTTCTGGACCC 3490
Qy       1034 ---IleCysIleThrAsnLysAsnProAlaAspPheThrThrIleSerAsnAspAsnGlu 1052
Db       3491 ATTGTGTGCAGCGTCAATAGAAACCCTGCTGATTTCACCATTCCTGAACCTGGGAATGTT 3550
Qy       1053 TyrMet 1054
Db       3551 TACATG 3556
```

RESULT 9
US-10-764-420-1635
; Sequence 1635, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; TITLE OF INVENTION: Possesses A Defined Biological Activity
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1635
; LENGTH: 7036
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-1635

Alignment Scores:

Pred. No.:	6.39e-10	Length:	7036
Score:	241.50	Matches:	266
Percent Similarity:	33.3%	Conservative:	162
Best Local Similarity:	20.7%	Mismatches:	428
Query Match:	4.4%	Indels:	432
DB:	9	Gaps:	60

US-09-828-068-2 (1-1057) x US-10-764-420-1635 (1-7036)

```
Qy        81 IlePheHisAspGlnLysLysCys---------------------------------- 88
             :::||||||:::   |||  |||
Db       515 GTTTTTCATAACTACAAAGCCTGCTCCACACCTGGATGGGGTTCATGTTGTTTTTGGACT 574

Qy        89 ------------------------------------AspGluHisLys-AlaSerSe 95
                                                 |||:::::: |||:::||
Db       575 GGTAATATCTGGTTTTGAAGTAATTGAACAGATTGAAAATCTGAAAACAGATGCTGCAAG 634

Qy        95 rSerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLe 115
             |  |||:::      |||      |||      ||||||                ||||||
Db       635 CAGACCTTAT------GCAGATGTCCGAGTTATTGACTGTGGGGTGCTGGCCACAAAGTT 688

Qy       115 uLysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSe 135
             |  |||  |||                |||      |||                |||
Db       689 G---ACAAAAGATGTTTTTGAGAAAAAAAGGAAGAAACCAACCTGTTCAGAAGGC----- 740

Qy       135 rAspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGl 155
                                                        |||      |||
Db       741 ---------------------------------------------TCGGACTCTTCTTC 754

Qy       155 nLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeu-- 174
              :::   |||   |||:::::||||||                 :::|||   |||:::
Db       755 CCGTTCCTCTTCCTCTTCAGAGTCCTCCTCAGAGAGTGAAGTTGAGCGAGAGACAATCAG 814

Qy       175 ----------------------ProLys---------------------------------- 176
                                    ||||||
Db       815 AAGGAGAAGACATAAGAGGAGGCCAAAAGTCAGACATGCTAAAAAGAGACGGAAAGAAAT 874

Qy       177 ----SerValGlnGluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAl 195
                 |||   :::|||          :::      :::|||   |||   :::
Db       875 GAGCAGTTCAGAAGAACCGAGGAGGAAGCGCACAGTAAGCCCTGAAGGTTATTCTGAGAG 934

Qy       195 aAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAs 215
              ::::::   |||                                    :::
Db       935 GAGTGATGTGAAT-----------------------------------GAAAAAAGATC 958

Qy       215 pValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluVa 235
             |||   :::|||          :::          |||                ::::||   ||
Db       959 AGTTGACTCAAACACTAAAAGAGAAAAGCCTGTTGTCCGCCCAGAAGAGATTCCTCCAGT 1018

Qy       235 lProGln---------------------------------IleThrTrpHisIleGluVa 244
             ||||:::                               ||||||          |||
Db      1019 TCCCGAGAACCGATTTTTACTTAGAAGAGATATGCCTGCTATCACTGTGGAGCCTGAACA 1078

Qy       244 lAsn--------------------GlyAlaAspGlnProProSer-------------- 252
              |||                   :::|||||   ||||||
Db      1079 GAACATTCCAGATGTTGCACCTGTTGTAAGTGATCAGAAACCCTCTGTATCAAAGTCTGG 1138

Qy       253 -----------------------------------ThrProLysLeuSerGluValVa 260
                                                 ||||||     |||
Db      1139 ACGGAAAATCAAAGGAAGAGGCACGATTCGCTATCACACACCTCCAAGGTCAAGATCCCA 1198

Qy       260 lLeuLysArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCy 280
                 :::      :::|||::::::   :::|||          |||
Db      1199 CTCTGAGTCCAAAGATGATGACAGCAGTGAAACCCCTCCTCACTGGAAGGAGGAGATGCA 1258

Qy       280 sAsnLeuThr------LysAspProAsnProMetSerGlyLysGluArgAspGlnValAl 298
                 |||         :::   |||:::              :::   |||::::::::
Db      1259 GAGACTGAGAGCCTACAGGCCCCCGAGCGGAGAGAAGTGGAGCAAAGGAGACAAGCTGAG 1318

Qy       298 aGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIl 318
             ::::   |||:::          |||   :::   :::|||   :::
```

```
Db      1319 TGACCCCTGTTCAAGCCGATGGGATGAAAGAAGCCTGTCCCAGAGATCCAGATCATGGTC 1378

Qy       318 eCysAsnGluProCysGluGluValValLeuLysArgSerSer---------------- 332
                |||          ::::::          |||   |||
Db      1379 CTATAATGGATACTATTCAGATCTTAGTACAGCGAGACACTCTGATGGTCACCATAAGAA 1438

Qy       333 -------LysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHis----- 348
                    |||   |||:::          ||||||      ||||||||||:::|||
Db      1439 ACACAGAAAGGAAAAGAAGTTTAAGCATAAAAAAAAAGCTAAAAAGCAGAAACATTGCAG 1498

Qy       349 ---------------SerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLe 363
                             |||:::|||              |||:::      :::
Db      1499 AAGACACAGACAGACAAAAAAGAGGAGAATAGTTATGCCTGATTTGGAACCCTCA----- 1553

Qy       363 uCysArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnVa 383
                |||  |||          |||:::  |||      :::      :::
Db      1554 -------AGATCTCCCACCCAC---CGAATGAAGTCCTCTTGTGTTAGAGAAAGGAGATC 1603

Qy       383 lGluAspSerArgSerAspGluValHis------ArgGluAsnAlaAlaAspProCysGl 401
                |||  |||          |||      :::  :::  :::  :::         ::
Db      1604 TCGTGCCTCCTCCTCCTCCTCTCATCACTCATCCAAGCGCGACTGGTCTAAATCAGACCA 1663

Qy       401 uAspAspArgSerThrIle-ProValProMetGluValSerMetAsp------------ 416
             :||||||  |||       |||:::|||  |||          |||
Db      1664 GGATGACGGGAGTGCTTCAACCCATTCCAGCAGAGACTCCTACAGATCCAAGTCTCATTC 1723

Qy       417 ----------------------------------IleProValSerAsnHisThrVal- 424
                                              :::      |||   |||   |||
Db      1724 ACGATCAGATTCTAGAGGGAGCTCTAGATCAAGGGCTGTGTCAAAGTCCTCATCTCGTTC 1783

Qy       425 -------------------------------GlyGluAspGly---------------- 428
                                            |||:::|||
Db      1784 TCTCAACAGATCAAAATCTAGATCTAGTTCCAGGTCAGGACCCCGAAGAACATCAATATC 1843

Qy       429 ----------------------LeuLysSerSerLysAsnLysThrLysArgLysTyrS 441
                                     :::|||   |||:::   ||||||:::
Db      1844 CCCCAAAAAACCTGCTCAGCTGAGTGAAAATAAGCCAGTT-AAGACAGAACCTTTAAGGC 1902

Qy       441 erAspValValAspAspGlySerSerLeuMet--------------------------- 451
                |||      :::|||:::   |||:::
Db      1903 CGTCAGTGCCACAGAATGGAAATGTGCTAGTGCAACCAGTGGCAGCAGAAAACATTCCTG 1962

Qy       452 -------------------------------AsnTrpLeuAsnGlyLysLys------- 458
                                             |||      |||:::|||
Db      1963 TAATACCATTGAGTGACAGCCCTCCCCCTTCTAGGTGGAAGCCTGGGCAGAAGCCCTGGA 2022

Qy       459 -------------------------LysArgThrGlySerValHisHisThrValAlaH 470
                                       :::      :::   |||
Db      2023 AGCCCTCTTACGAGCGAATTCAGGAGATGAAAGCTAAAACAACCCACTTGCTGCCTGTCC 2082

Qy       470 isProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHisAspA 490
                  :::|||:::|||   |||   |||   ::::::|||:::   |||
Db      2083 AAAGCACATACAGCTTAACAAATATTAAAGCAACCGTGAGTTCATCATCTTATCACAAAA 2142

Qy       490 spGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCys---- 508
              |||      :::|||:::                 :::|||
Db      2143 GAGAAAAACCTTCAGAAAGTGATGGGAGTGCTTATTCAAAGTACAGTGATAGAAGTTCTG 2202

Qy       509 ----------GlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysT 525
                        :::      |||:::   |||      |||   ||||||:::     |
Db      2203 GAAGCTCAGGAAGGTCGGGGAGCAAGTCTTCTAGGAGCAGGTCATCCTCCAGGTCCTACA 2262

Qy       525 hrAlaGlyLeuSerLys-------------------------------------GlyL 532
             ||      |||:::                                        :
Db      2263 CAAGGTCAAGGTCAAGAAGTCTCCCTACTTCACGCTCACTCTCTAGGTCTCCATCATCTA 2322

Qy       532 ysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnI 552
             :::::||||||                |||   :::   :::|||:::
Db      2323 GGTCTCACTCACCAAATAAGTACAGTGATGGTTCCCAGCACAGTAGGTCATCTTCATATA 2382

Qy       552 leHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsn---------SerV 569
                |||   |||   :::   :::      |||          |||          ||||
Db      2383 CTTCTGTTAGCAGTGATGATGGAAGACGAGCCATGTTTAGATCCAACAGGAAAAAAAGTG 2442
```

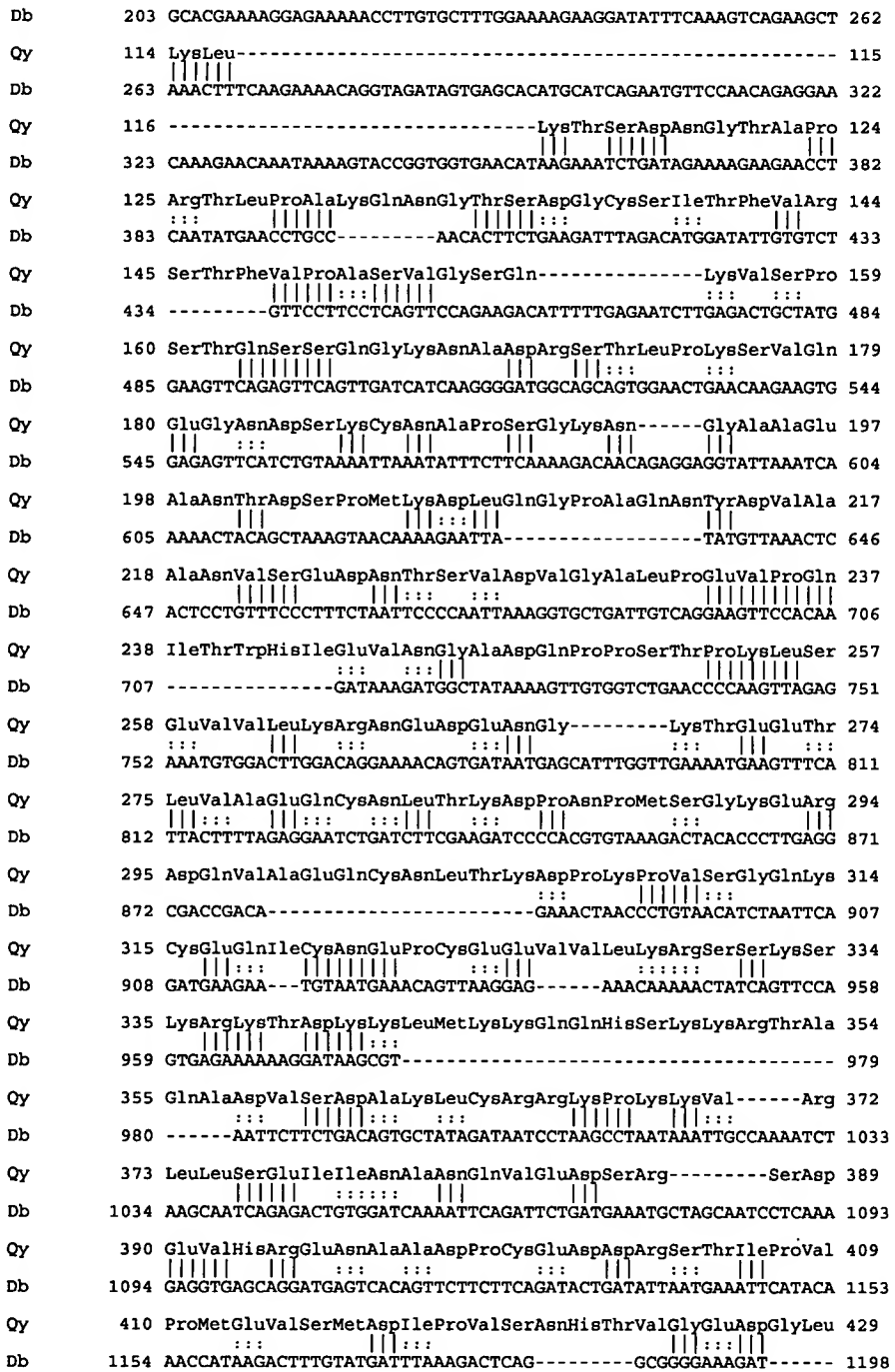
```
Qy        569 alLeuSerHisSer---------------------------------------------- 573
              ||  ||||||
Db       2443 TCACTTCACATAAAAGACATCGCAGCAACTCTGAAAAGACACTTCACAGTAAATATGTCA 2502

Qy        574 -------------------------AlaLysValSerProAlaGluHisAspIleGlnI 585
                                       |||  |||  :::      :::
Db       2503 GAGGCAGAGAGAAATCCTCACGTCACAGAAAGTATAGTGAAAGTAGATCATCTTTAGATT 2562

Qy        585 leMetSerAsp---------LeuHisGluGln------SerLeuProLysLysLysLysL 600
              ||||||           |||  |||      |||  |||:::|||:::|||:
Db       2563 ACACTTCAGACAGTGACCAGTCACATGTTCAAGTATACTCAGCCCCAGAGAAGGAGAAGC 2622

Qy        600 ysGlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspI 620
              ::  |||:::|||        :::||||||
Db       2623 AGGGAAAAGTGGAAGCATTGAATGATAAGCAG---------------------------- 2654

Qy        620 leValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysS 640
                          |||  :::  |||  :::          |||  :::|||
Db       2655 -------------GGGAAAGGCAGAGAAGAAGGAAAACCCAAGCCTGAATGGGAATGTC 2700

Qy        640 erAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaA 660
                    :::      |||:::     :::     |||||  |||
Db       2701 CTCGTTCTAAAAAAGAGAACTCCGAAGATCACTCTAGAGATGACAGTGTG---------- 2750

Qy        660 laLysAspGlySerAspTyrAlaSerSerValPheAsp------ThrAsnSerGlnGln- 677
              |||  :::  |||  |||  :::|||      :::|||||:::|||
Db       2751 --TCCAAAGGGAAGAATTGTGCGGGGAGTAAATGGGATTCGGAATCAAACTCAGAACAAG 2808

Qy        678 -----------------LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyH 692
                              ||||||         :::   :::::||||||  :::|||
Db       2809 ATGTGACTAAGAGCAGGAAAAGTGATCCCCGGAGAGGTTCAGAAAAGGAG---GAGGGTG 2865

Qy        692 isLeuAla-LeuThrThrGlnGlu---------------------SerProHisProGln 704
                :::    |||      :::                      |||:::||||||
Db       2866 AAGCCTCTTCAGACTCCGAGTCAGAAGTTGGTCAGAGTCACATCAAAGCCAAACCCCCAG 2925

Qy        705 AsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValThrIle 724
                  ||||||  |||                  :::  :::
Db       2926 ------CAAAGCCTCCAA----------------------CAAGCACTTTTCTGCCCG 2955

Qy        725 AlaAlaSerSerPro------LeuPheSerHisHisAspAspGlnTyrIleAlaGluAla 742
              |||||||:::  |||      |||      |||                      |||
Db       2956 GCAGCGACGGTGCCTGGAAGTCTAGGAGACCACAGTCTTCAGCCTCTGAGTCAGAGAGCT 3015

Qy        743 ProThrGluHisTrpGly---------------------ArgLysAspAlaLysLysLeu 755
              |||      ||||||                    |||:::  :::::::::
Db       3016 CCTGCTCCAACTTGGGGAACATTAGAGGAGAGCCCCAGAAGCAAAAACACTCAAAGGATG 3075

Qy        756 ThrTrp-----------------------------GluGlnPheLysAlaThrThrArg 765
                                              |||          |||:::
Db       3076 ATCTTAAGGGGGATCACACAAAAAGGGCAAGAGAGAAGTCAAAAGCTAAAAAAGACAAAA 3135

Qy        766 AsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGly---------IleGlnAla 782
              |||:::                |||::::::       |||
Db       3136 AACACAAGGCTCCAAAAC---GGAAGCAAGCTTTCCACTGGCAACCTCCACTCGAGTTTG 3192

Qy        783 ValAspLeuThrSerThrHisValMetGlySer-------------------------- 793
              |||  :::             |||||||||
Db       3193 GTGACGATGAGGAGGAGGAGATGAATGGGAAGCAAGTTACACAGGACCCAAAAGAGAAAA 3252

Qy        794 --------------------------------SerSerAsnTyrAlaSerArgGlnPro 802
                                             ::::::      ||||||:::|||
Db       3253 GGCATGTCTCTGAGAAGTGTGAAGCTGTGAAAGACGGCATTCCAAACGTCGAGAAAACCT 3312

Qy        803 ValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsn 822
              |||:::                        :::|||||||  |||||||:::  |||
Db       3313 GTGATG----------------------AAGGCAGTTCTCCAAGTAAACCCAAGAAGG 3348

Qy        823 PheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGly 842
                                                  |||      |||  :::
Db       3349 GTA-------------------------------CTTTAGAGCAGGACCCACTTGCAG 3375

Qy        843 GlnValValLeuTyrPro-------------------------LysGluSerMetPro 853
              :::|||  :::  |||                          :::  :::  |||
Db       3376 AGGGTGGACATGATCCCAGCTCTTGTCCTGCACCTCTGAAAGTGGAGGACAACACGGCCA 3435

Qy        854 AlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSer-PheProAsnTyr-- 872
              |||  ||||||                |||||||||||  |||:::|||  :::
Db       3436 GCTCTCCACCTAGCG----------CCCAGCACCTTGAAGAGCATGGCCCAGGTGGAG 3483

Qy        873 ----GlyThrSerSerArgAsnGlnMetGluSerGlnLeuHis-AsnSerGlnTyrAlaH 891
                  |||  :::|||                    ||||||      :::
Db       3484 GGGAGGACGTGCTTCAGACAGATGACAACATGGAGATTTGCACGCCTGATAGG------- 3536

Qy        891 isAsnGlnTyrLysGlySerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysIleP 911
                                  ||||||  :::  |||
Db       3537 -------------------ACTTCCCCTGCAAAGGGAGAG------------------- 3557

Qy        911 roLeuThrPheGluAspLeuSerArgHisGlnLeuHisAspLeuHisArgProLeuArgP 931
                        |||:::  |||:::|||                  :::   |
Db       3558 -----GTGGTGTCCCCTTTAGCAAACCACAGGCTAGACAGCCCAGAGGTGAACATTATTC 3612

Qy        931 ro--------------------HisProArgValGlyValLeuGlySerLeuLeuGlnL 944
              ||                  |||||||||  |||      |||      |||:
Db       3613 CAGAGCAGGATGAGTGTATGGCACATCCTAGAGCAGGA------GGAGAA------CAAG 3660

Qy        944 ysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValS 964
              ::          :::  |||:::      |||  :::||||||  |||       |
Db       3661 AGAGTAGCATGTCTGAAAGCAAGACCTTGGGT---GAAAGTGGGGTTAAACAGGACAGCT 3717

Qy        964 erThrGlyIleThrSer------HisGlnMetAsnArgLysGluHisPheGluAlaLeuA 982
              |||||   :::||||||                :::|||||||      |||
Db       3718 CTACCAGTGTGACCAGTCCTGTAGAAACTTCTGGAAAGAAGGAGGGGGCTGAGAAGAGCC 3777

Qy        982 snSerGlyMetPheSerAlaLysTrpAsnAlaLeuGln---LeuGlySerValSer---- 999
                  :::   :::  ||||||     ||||||  :::|||::::::|||
Db       3778 AAATGAACCTC---ACAGATAAGTGGAAGCCATTGCAAGGTGTAGGGAATCTGTCAGTGT 3834

Qy       1000 -------------SerSerAlaAspPheLeuSerAlaArgAsnSerIleAlaGlnSerT 1015
                           |||||||||      :::  |||  :::::::::  :::
Db       3835 CTACTGCAACCACATCCAGTGCTCTGGATGTGAAGGCATTATCTACTGTGCCTGAAGTGA 3894

Qy       1015 rpThrArgGly---------------LysGlyLys---MetValHisProLeuAspArgP 1029
              :::|||                |||  |||      |||  |||       |
Db       3895 AACCACAAGGCTTGAGGATAGAAATCAAAAGCAAAAATAAGGTTCGGCCTGGGTCTCTCT 3954

Qy       1029 he------ValArgGlnAspIleCysIleThrAsnLysAsnProAlaAspPheThrThrI 1047
              ||      ||||||:::         :::   |||:::   |||   :::     :::
Db       3955 TTGATGAAGTAAGAAAGACGGCACGCCTA---AATCGAAGGCCACGGAATCAAGAGAGTT 4011

Qy       1047 leSerAsnAsp 1050
                |||:::|||
Db       4012 CCAGTGATGAT 4022
```

```
RESULT 10
US-10-450-763-8960
; Sequence 8960, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 8960
;   LENGTH: 5286
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (49)..(312)
;   OTHER INFORMATION: 86% homologous to Homo sapiens line-1 reverse
;   OTHER INFORMATION: transcriptase,accession number U83303,Smith-Waterman Score=377.
US-10-450-763-8960

Alignment Scores:
Pred. No.:              1.63e-08        Length:         5286
Score:                  223.50          Matches:        210
Percent Similarity:     34.2%           Conservative:   173
Best Local Similarity:  18.8%           Mismatches:     412
Query Match:            4.0%            Indels:         325
DB:                     9               Gaps:           52

US-09-828-068-2 (1-1057) x US-10-450-763-8960 (1-5286)

Qy         83 HisAspGln---------LysLysCysAspGluHisLysAlaSerSerSerProPheSer 99
              |||||:::           ||||||     |||          :::
Db       2236 CATGATGAAGTGGTTGATGTCAAATGCCATTCACACTCTCCTTTTAGGAATGAAAGAGGA 2295

Qy        100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119
                 |||   |||                |||::::::|||||      ::::::
Db       2296 AAAGGAAAAATAAGGCATCAT------ATATCCTGTATTGAAAAGTTAAGCAAAACAGAA 2349

Qy        120 AsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSer 139
              :::  :::  |||                        |||||||||      |||
Db       2350 AGTATATCAGTACCC------------------------ACCAGTGATCACAGGAGC 2382

Qy        140 IleThrPheValArgSerThrPheValProAlaSerValGlySerGlnLysValSer--- 158
              :::                    :::   |||:::      :::   |||||||||
Db       2383 CTC--------------------ATTGAAGCAAATCAAAGCAATTCCAAAGTTTCTGAA 2421

Qy        159 ------------------------ProSerThrGlnSerSer---------GlnGlyLys 167
                                       |||         ||||||            |||:::
Db       2422 CTTGACACAATTTATTGTACCTTGCCAAGAAAATCAAGCAGTTTTCTCATACATGGCAGG 2481

Qy        168 AsnAlaAspArgSerThrLeuProLysSerValGlnGluGly------------------ 181
                  :::       :::      |||::::::  |||
Db       2482 CAGTCAGGAAGTAAAATAATGGCTGCTTCATTGAGGAATGGGCCACCTCCCTTCCAAATC 2541

Qy        182 ------------------------------------AsnAspSerLysCysAsnAla 188
                                                  |||          ::::::
Db       2542 AAAAATAATGTGGAAGATGCAATGGGGAACTATATGTTAAACAAATTTAGTCCCAGTTCT 2601

Qy        189 ProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln 208
              |||         |||      ::::::     :::|||||                |||:::
Db       2602 CCTGAGTCAGCGAATGAATGTTCCAAAGTCCTTTCAGACTCAGCC---------CTGGAA 2652

Qy        209 GlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAsp 228
                 |||          :::          |||      :::|||||||||
Db       2653 GCACCTGAAGCCACAGAGAGAATGACAAATGTAAAAAGCAGTGGATCTACTTCCGTTAGA 2712

Qy        229 ValGlyAlaLeuPro--------------------GluValProGlnIleThrTrpHis 241
                 |||   ||||||                   |||                 |||
Db       2713 AAAGGACCACTTCCATTCCTCATCAACAGGGCTATGTCATGTCCCTCAGGGGAGCCACAT 2772

Qy        242 IleGlu-----------------------------ValAsnGlyAlaAspGlnProPro 251
                                                 :::|||   |||
Db       2773 GCCTCAACTGGAAGAGAAGGAAGAAAAAAGCCATTGACCTCAGGCATGGATGCTTCTGAG 2832

Qy        252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
                 ||||||:::     |||   ::::::         |||   ::::::
Db       2833 CTAACACCAAGGGCTTGGGAGAGAATCATTAGCCCTGTGGAAAGTGACTCATCT------ 2886

Qy        272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLys--------------------- 284
                                       |||:::||||||||
Db       2887 --------------GTTAGAGATTGTTCTTTAACCAAAAGACAACACCAAAAGGAAAAC 2931

Qy        285 --------------AspProAsnProMetSerGlyLysGluArgAspGlnValAlaGlu 299
                          :::         :::|||||||||||||||   :::
Db       2932 TTCCAAGAATACACTGAGAAAGAGGACTTGTCAGGAAAAGAACGTGGGAAAACATTACAT 2991

Qy        300 GlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCys 319
              :::          |||                 |||||||||   :::          |||
Db       2992 AAAGTTAAG---ACGACTAGTACGTTTTCTGTTTCTGGTGATGAAGATAATGTAAAATGT 3048

Qy        320 AsnGluPro-CysGluGluValValLeuLysArgSerSerLysSerLys---------- 335
                 |||       :::   :::   |||   |||   :::   :::|||
Db       3049 CTGGAGGTGGTCTCAATATATTACACTCTACCGAGGAAACCCAGCAAAAAATTCTGTAAC 3108

Qy        335 ---------------------------------------------------------- 335

Db       3109 CTCCTTCAACAGTATACACAAAATACTAATTTACTTATAGAATCACCTCAAGTGGAGACT 3168

Qy        336 ----------------------ArgLysThrAspLysLysLeuMet------LysLysGl 346
                                     :::||||||::::::   ::::::          ||
Db       3169 GAAACATTTCCTAACGCTTTAGAAAAAGACAAACAGAATTATTCTACACGAGAGCAGTCA 3228

Qy        346 nGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgAr 366
              |:::|||      :::      :::          :::  |||  |||||
Db       3229 GGAACACCTTCATGTGAAAATCTAAAGATGT---------CCGTCAACTCTGATCAGACG 3279

Qy        366 gLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSe 386
              |  |||   |||:::   |||   |||:::      |||          :::      ::
Db       3280 CTCACCACTGAAAATATGACTGCCTTCCGATTATCA-AATAGGGGGCCCCTAGCTCCTAC 3338

Qy        386 rArgSerAspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerTh 406
              :      :::              |||   |||
Db       3339 ATTACAGGAAATGGCTTCTGTTGAGGCAGCT---------------------------- 3369

Qy        406 rIleProValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGl 426
               :::   :::|||   |||   |||          :::   :::      :::   ::
Db       3370 -GTTTCTCTTCCTGAAGAGGAATCTAAAGCTAGAGAGATTTTTTCAGATAATTTAGCTAA 3428

Qy        426 uAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAs 446
              :      |||      |||:::||||||   :::|||
Db       3429 AACACCTCTAGGTGATTCAGAAAACAAGAAGGAAAGA---------------------- 3465

Qy        446 pGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHi 466
                                          |||||||||   ::::::   ::::::|||
Db       3466 ---------------------------GGCAAAAAGTTGCAAAGTGAAACCCTGCATAC 3497

Qy        466 sThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSerTh 486
               ::::::                |||:::|||   :::|||   :::
Db       3498 TTCATTGATGCTTCAGAGAAAAAATGTATCCGAAGAAAAA---TCTGAAAATTGTCAACA 3554

Qy        486 rGlnHisAspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAs 506
                  :::      |||          :::   |||          :::   :::   ::
Db       3555 ATCCATTAATTCAAGTAACAGTGGTCCCTCTAGTCTTCCAGCT---CTTTCAGAAGTTAA 3611

Qy        506 pValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThr-- 525
               ::::          :::      |||:::      |||:::      |||:::
Db       3612 TATTGGAAATTCCCAAACTAGAAGAAGTTCTTGGGAGTGTACAGGGAGTGGTAGAGCCAT 3671

Qy        526 ------------AlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTy 541
                         :::|||          |||   |||::::::   :::
Db       3672 TCCATTTACTGGAAGTGGCAAGTGTCCCCAGAAAGATCACACATCCACAGCTGTAGGTGA 3731

Qy        541 rGly------GlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAs 559
              |||      |||          |||   |||:::
Db       3732 TGGCTCCAGTGGATCACAGCCTAGGGAAGGCAGAGGG-----------GACATTGGAAC 3779

Qy        559 pGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSerAlaLys---------- 575
                 ||||||      |||   |||       ||||||||||||
Db       3780 CAACTGCCAAAAAATGACT---AATAAAACACTTTCTCACTCAGAGAGTCAAGTCTTTGC 3836

Qy        576 -ValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGlu---------- 591
               ::::::||||||   |||   :::|||:::      :::
Db       3837 CCTTACTCCAGCATTGCATAAACTACAGCTTGGTGAGGAGACTCAGTCAGATGAACCAAA 3896

Qy        592 ------------GlnSerLeuProLysLys------LysLysLysGlnLysLeuGluVa 605
                          ||||||   |||::::::      ::::::   |||:::      ::
Db       3897 CTTAGAGAGTCTGCAGTCTGAACCAAGAGAATTACCTCAAAGAAGTCAGGAGGCAAATAT 3956
```

```
Qy       605 lThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeuAl 625
Db      3957 GACAGAGAGCAGGAAG----------------------------------------GC 3974

Qy       625 aLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIl 645
Db      3975 TGAAGATGAAATGCAGAAGTCAGCTTGGGATCAACCTTCACTTCCTGAAGGAAACAAA-- 4032

Qy       645 eGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAs 665
Db      4033 ----AATAAAACCAACTTGGATGAC---------CTAGTAAAGGGGGAAAATAGATCTTC 4079

Qy       665 pTyr--------------------------AlaSerSerValPheAspThrAsn----- 674
Db      4080 AGTTAAACACAGATTGGCAGCCATGTCTAAAGCCAGCAGAAAATTCCCAGCTAAAGATGT 4139

Qy       675 -SerGlnGlnLysSerLeuAlaSer---------GlnSerThrGlnLysGluLeuGlnGl 691
Db      4140 AAGCCCCAGAAGACATGTAGCTACTATCTTCCCCCAAAGTGGAAGCAGATCTGGCTTTGA 4199

Qy       691 yHisLeuAlaLeuThrThrGlnGlu-----------------SerProHisProGlnAs 705
Db      4200 CCATTTATCTCTTGGCACAGTGGAGTGCAACCCACTGTTCCCTGAGCCTACTCCAAAATC 4259

Qy       705 nPheGlnSerThrGlnGluGlnGln---------------ThrHisLeuArgMetGluGl 720
Db      4260 TGCAGAGTCCATTGGCGAAAGCAGGTTGAGTGAGAACGGAAAGCATGTGAAGAAATCCGA 4319

Qy       720 u-----------------------------------------MetValThrIleAl 725
Db      4320 GAACCTTCTCCCCATTACTGTACTACCCAACAGAGAACCTTCTACACACGTCAGCAACCA 4379

Qy       725 aAlaSerSerProLeuPheSerHisHisAspAspGlnTyr------IleAlaGluAlaPr 743
Db      4380 GAAGTCTAACAGCATTTCACAACGACATCAGAATGAGTTTAAAAACGTCTCAGAATCACC 4439

Qy       743 oThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaTh 763
Db      4440 ATCAAAGCATGAGAATTCTAAAGAC----------------------------GTCAC 4469

Qy       763 rThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaVa 783
Db      4470 AGCAGCTCAGAATTTAGTAAGAGAATCAGGAGCCCCA--------------TCACCCAT 4514

Qy       783 lAspLeuThrSerThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProVa 803
Db      4515 CACATTCACCAGCCTCAGGGAAGCAGAATTCTCTGACAATCAGAGGAGGCTGAGCCCTCC 4574

Qy       803 lIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPh 823
Db      4575 TTTT---CCACTGGAGCCT---GCACAGAAA---------------------------- 4599

Qy       823 eProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAsnAlaGlyGl 843
Db      4600 ----TCTAGAGTAAGCAGTCCACTGGCCAGTTTTCTGCAGCAACAAAGGAGTGCTTCATC 4655

Qy       843 nValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMetMetAspPr 863
Db      4656 TCTG--------------GAGTGGGAACCTGAGCCACACCTCTATCGT---------TC 4691

Qy       863 oSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMetGluSerGl 883
Db      4692 AAAGAGTTTAAAAAGCATTAATGTTCATGGCGATCTACTACGAAAAAGC----------- 4740

Qy       883 nLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerThrSerTyrGl 903
Db      4741 ----CATCCTCCAAAAGTCAGGGAGCGCCATTTTTCTGAAAGCACTTCTATT-------- 4788

Qy       903 ySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHisGlnLeuHi 923
Db      4789 -GACAATGCCCTGAGTCGACTGACCCTTGGGAATGAATTCTCTGTCAACAATGGGTACAG 4847

Qy       923 sAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySerLeuLeuGl 943
```

```
Db      4848 TCGAAGATTCAGATCTTTTTCTGAACTCCCCTCCTGTGAT------GGAAAT-------- 4893

Qy       943 nLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyVa 963
Db      4894 ----------GAAAGTTGGGCTTATCGCAGCGGGACAAAAACAGGTCCCAGGTCTGCAAT 4943

Qy       963 lSerThrGlyIleThrSerHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSe 983
Db      4944 ATCTATA-----------------------------------TACAGACCTATCGACTA 4967

Qy       983 rGlyMetPheSerAlaLysTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAs 1003
Db      4968 TGGGATCTTTGGGAAAGAACAACAGTTAGCTTTCTTAGAAAATGTAAAGAGGTCA----- 5022

Qy      1003 pPheLeuSerAlaArgAsnSerIleAlaGlnSerTrpThrArgGlyLysGlyLysMetVa 1023
Db      5023 ----CTTACACAAGGAAGATTATGGAAACCAAGTTTTCTTAAGAACCCTGGC-------- 5070

Qy      1023 lHisProLeuAspArgPheValArgGlnAspIleCysIleThrAsnLysAsnPro 1041
Db      5071 ---------------TTCCTAAAAGATGATTTG------------AGGAACCCT 5097

RESULT 11
US-10-029-386-22571
; Sequence 22571, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22571
;   LENGTH: 2612
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO Z84487.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.74
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
;   OTHER INFORMATION: NT HIT: gi14757764, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AW769265.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P46100, EVALUE 0.00e+00
US-10-029-386-22571

Alignment Scores:
Pred. No.:              6.91e-09        Length:         2612
Score:                  223.00          Matches:        172
Percent Similarity:     34.3%           Conservative:   108
Best Local Similarity:  21.1%           Mismatches:     312
Query Match:            4.0%            Indels:         225
DB:                     6               Gaps:           33

US-09-828-068-2 (1-1057) x US-10-029-386-22571 (1-2612)

Qy         62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
Db         86 AAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTA 145

Qy         82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
Db        146 ---AACAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTGAAACAAAA 202

Qy        101 --------------------AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
```



```
Qy        731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TrpGlyArgLy 750
Db       2106 -----------------CAGAGGATAAAAAGAGTAAGAATGGAGCATATGGTAGAGAGA 2147

Qy        750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db       2148 AGAAAAGGTGCAAGTTGCTTGGAAAGAGTTCAAGGAAGAGACAAGA 2193

RESULT 12
US-10-956-157-4592
; Sequence 4592, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4592
;   LENGTH: 10217
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-4592

Alignment Scores:
Pred. No.:              4.41e-08        Length:         10217
Score:                  223.00          Matches:        172
Percent Similarity:     34.3%           Conservative:   108
Best Local Similarity:  21.1%           Mismatches:     312
Query Match:            4.0%            Indels:         225
DB:                     9               Gaps:           33

US-09-828-068-2 (1-1057) x US-10-956-157-4592 (1-10217)

Qy         62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
Db       1311 AAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTA 1370

Qy         82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
Db       1371 ---AACAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTGAAACAAAA 1427

Qy        101 ----------------------AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
Db       1428 GCACGAAAAGGAGAAAAACCTTGTGCTTTGGAAAAGAAGGATATTTCAAAGTCAGAAGCT 1487

Qy        114 LysLeu---------------------------------------------------- 115
Db       1488 AAACTTTCAAGAAAACAGGTAGATAGTGAGCACATGCATCAGAATGTTCCAACAGAGGAA 1547

Qy        116 --------------------------------LysThrSerAspAsnGlyThrAlaPro 124
Db       1548 CAAAGAACAAATAAAAGTACCGGTGGTGAACATAAGAAATCTGATAGAAAAGAAGAACCT 1607

Qy        125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
Db       1608 CAATATGAACCTGCC---------AACACTTCTGAAGATTTAGACATGGATATTGTGTCT 1658

Qy        145 SerThrPheValProAlaSerValGlySerGln--------------LysValSerPro 159
Db       1659 ---------GTTCCTTCCTCAGTTCCAGAAGACATTTTTGAGAATCTTGAGACTGCTATG 1709

Qy        160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
Db       1710 GAAGTTCAGAGTTCAGTTGATCATCAAGGGGATGGCAGCAGTGGAACTGAACAAGAAGTG 1769

Qy        180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn------GlyAlaAlaGlu 197
Db       1770 GAGAGTTCATCTGTAAAATTAAATATTTCTTCAAAAGACAACAGAGGAGGTATTAAATCA 1829

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db       1830 AAAACTACAGCTAAAGTAACAAAAGAATTA-----------------TATGTTAAACTC 1871

Qy        218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
Db       1872 ACTCCTGTTTCCCTTTCTAATTCCCCAATTAAAGGTGCTGATTGTCAGGAAGTTCCACAA 1931

Qy        238 IleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSer 257
Db       1932 --------------GATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAG 1976

Qy        258 GluValValLeuLysArgAsnGluAspGluAsnGly---------LysThrGluGluThr 274
Db       1977 AAATGTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCA 2036

Qy        275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
Db       2037 TTACTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTTGAGG 2096

Qy        295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314
Db       2097 CGACCGACA-----------------------GAAACTAACCCTGTAACATCTAATTCA 2132

Qy        315 CysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSer 334
Db       2133 GATGAAGAA---TGTAATGAAACAGTTAAGGAG------AAACAAAAACTATCAGTTCCA 2183

Qy        335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354
Db       2184 GTGAGAAAAAAGGATAAGCGT------------------------------------ 2204

Qy        355 GlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal------Arg 372
Db       2205 ------AATTCTTCTGACAGTGCTATAGATAATCCTAAGCCTAATAAATTGCCAAAATCT 2258

Qy        373 LeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg---------SerAsp 389
Db       2259 AAGCAATCAGAGACTGTGGATCAAAATTCAGATTCTGATGAAATGCTAGCAATCCTCAAA 2318

Qy        390 GluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProVal 409
Db       2319 GAGGTGAGCAGGATGAGTCACAGTTCTTCTTCAGATACTGATATTAATGAAATTCATACA 2378

Qy        410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429
Db       2379 AACCATAAGACTTTGTATGATTTAAAGACTCAG---------GCGGGGAAAGAT------ 2423

Qy        430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp--------- 446
Db       2424 ------GATAAAGGAAAAAGGAAACGAAAAAGTTCTACATCTGGCTCAGATTTTGATACT 2477

Qy        447 ------GlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg----------- 460
Db       2478 AAAAAGGGCAAATCAGCTAAGAGCTCTATAATTTCTAAAAAGAAACGACAAACCCAGTCT 2537

Qy        461 --------------------------ThrGlySerValHisHisThrValAlaHisPro 471
Db       2538 GAGTCTTCTAATTATGACTCAGAATTAGAAAAAGAGATAAAGAGCATGAGTAAAATTGGT 2597

Qy        472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSer---ThrGlnHisAspAsp 490
Db       2598 GCTGCCAGAACCACCAAAAAAAGAATTCCAAATACAAAAGATTTTGACTCTTCTGAAGAT 2657

Qy        491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr---AspValCysGln 509
Db       2658 GAGAAACACAGCAAAAAAGGAATGGATAATCAAGGGCACAAAAATTTGAAGACCTCACAA 2717

Qy        510 HisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529
Db       2718 GAAGGATCATCTGATGATGCTGAAAGA---AAACAAGAGAGAGAGACTTTCTCTTCAGCA 2774
```

```
Qy        530 LysGly-------------------------------------------------Lys 532
              :::|||                                                 |||
Db       2775 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTAAGAGATCGACTTCCTAAGAAG 2834

Qy        533 ThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIle 552
                :::::|||||||||         |||           :::|||:::
Db       2835 CAGCAAGCAAGTGCTTCCACT------GATGGTGTCGATAAGCTTTCTGGGAAAGAGCAG 2888

Qy        553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHis 572
                  |||  |||  :::        ||||||:::                  |||
Db       2889 AGTTTTACTTCTTTGGAAGTTAGAAAAGTTGCTGAAACTAAAGAAAAG------AGCAAG 2942

Qy        573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592
                |||          :::             :::|||||:::   |||:::
Db       2943 CATCTCAAAACCAAAACATGTAAAAAAGTACAGGATGGCTTATCTGATATTGCAGAGAAA 3002

Qy        593 SerLeuProLysLysLys----------------------LysLysGlnLysLeuGlu 604
                |||  |||  :::                         |||||||||      :::
Db       3003 TTCCTAAAGAAAGACCAGAGCGATGAAACTTCTGAAGATGATAAAAAGCAGAGCAAAAAG 3062

Qy        605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
                |||  ||||||:::                  |||  |||      :::  :::
Db       3063 GGAACTGAAGAAAAAAAGAAA--------------CCTTCAGACTTTAAGAAAAAAGTA 3107

Qy        625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp--------- 641
                |||         |||:::|||      :::   :::|||   ::::::
Db       3108 ATTAAA------ATGGAACAACAGTATGAATCTTCATCTGATGGCACTGAAAAGTTACCT 3161

Qy        642 --------------------------------IleAsnArgIleGlnSerLysThrThr 650
                                              |||   :::|||:::::   ||||||
Db       3162 GAGCGAGAAGAAATTTGTCATTTTCCTAAGGGCATAAAACAAATTAAGAATGGAACAACT 3221

Qy        651 AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVal 670
                                             ||||||      :::      :::
Db       3222 --------------------------------GATGGAGAAAAGAAAAGTAAAAAAATA 3248

Qy        671 PheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690
                |||      |||::::::|||         |||         :::|||     |||
Db       3249 AGAGATAAAACTTCTAAAAAGAAGGATGAATTATCTGATTATGCTGAGAAGTCAA--CAG 3306

Qy        691 GlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGln 710
              |||
Db       3307 GGA---------------------------------------------------------- 3309

Qy        711 GluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeu 730
                                     |||   :::||||||:::
Db       3310 ---------------------AAGGAGATAGTTGTGACTCTT------------------ 3330

Qy        731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TrpGlyArgLy 750
                                  |||   |||          ||||||  |||   |||::
Db       3331 -----------------CAGAGGATAAAAAGAGTAAGAATGGAGCATATGGTAGAGAGA 3372

Qy        750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
              ::::      ::::::   |||::::::||||||      ||||||
Db       3373 AGAAAAGGTGCAAGTTGCTTGGAAAGAGTTCAAGGAAGAGACAAGA 3418

RESULT 13
US-10-956-157-426
; Sequence 426, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 426
;   LENGTH: 10330
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-426

Alignment Scores:
Pred. No.:              4.47e-08        Length:         10330
Score:                  223.00          Matches:        172
Percent Similarity:     34.3%           Conservative:   108
Best Local Similarity:  21.1%           Mismatches:     312
Query Match:            4.0%            Indels:         225
DB:                     9               Gaps:           33

US-09-828-068-2 (1-1057) x US-10-956-157-426 (1-10330)

Qy         62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
              :::   ::::::||||||   :::      :::   :::|||   ::::::      :::
Db       1425 AAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTA 1484

Qy         82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
                 :::  :::|||          |||||||||          :::   |||
Db       1485 ---AACAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTGAAACAAAA 1541

Qy        101 ----------------------AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
                                     :::   |||  ||||||
Db       1542 GCACGAAAAGGAGAAAAACCTTGTGCTTTGGAAAAGAAGGATATTTCAAAGTCAGAAGCT 1601

Qy        114 LysLeu---------------------------------------------------- 115
              ||||||
Db       1602 AAACTTTCAAGAAAACAGGTAGATAGTGAGCACATGCATCAGAATGTTCCAACAGAGGAA 1661

Qy        116 --------------------------------LysThrSerAspAsnGlyThrAlaPro 124
                                             |||  ||||||      |||
Db       1662 CAAAGAACAAATAAAAGTACCGGTGGTGAACATAAGAAATCTGATAGAAAAGAAGAACCT 1721

Qy        125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
                :::      ||||||          ||||||:::        :::      |||
Db       1722 CAATATGAACCTGCC---------AACACTTCTGAAGATTTAGACATGGATATTGTGTCT 1772

Qy        145 SerThrPheValProAlaSerValGlySerGln---------------LysValSerPro 159
                  ||||||:::||||||                           :::   :::
Db       1773 ---------GTTCCTTCCTCAGTTCCAGAAGACATTTTTGAGAATCTTGAGACTGCTATG 1823

Qy        160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
                  |||||||||          |||   |||:::      :::
Db       1824 GAAGTTCAGAGTTCAGTTGATCATCAAGGGGATGGCAGCAGTGGAACTGAACAAGAAGTG 1883

Qy        180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn------GlyAlaAlaGlu 197
                |||  :::      |||  |||      |||       |||       |||
Db       1884 GAGAGTTCATCTGTAAAATTAAATATTTCTTCAAAAGACAACAGAGGAGGTATTAAATCA 1943

Qy        198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
                  |||              |||:::|||                  |||
Db       1944 AAAACTACAGCTAAAGTAACAAAAGAATTA-----------------TATGTTAAACTC 1985

Qy        218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
                |||||      |||:::   :::                  ||||||||||||
Db       1986 ACTCCTGTTTCCCTTTCTAATTCCCCAATTAAAGGTGCTGATTGTCAGGAAGTTCCACAA 2045

Qy        238 IleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSer 257
                           :::   :::|||                  |||||||||
Db       2046 --------------GATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAG 2090

Qy        258 GluValValLeuLysArgAsnGluAspGluAsnGly---------LysThrGluGluThr 274
                :::      |||   :::         :::|||             :::   |||   :::
Db       2091 AAATGTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCA 2150

Qy        275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
                |||:::   |||:::   :::|||   :::   |||         :::         |||
Db       2151 TTACTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTTGAGG 2210

Qy        295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314
                                           :::      ||||||:::
Db       2211 CGACCGACA----------------------GAAACTAACCCTGTAACATCTAATTCA 2246
```

```
Qy       315 CysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSer 334
Db      2247 GATGAAGAA---TGTAATGAAACAGTTAAGGAG------AAACAAAAACTATCAGTTCCA 2297

Qy       335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354
Db      2298 GTGAGAAAAAAGGATAAGCGT------------------------------------- 2318

Qy       355 GlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal------Arg 372
Db      2319 ------AATTCTTCTGACAGTGCTATAGATAATCCTAAGCCTAATAAATTGCCAAAATCT 2372

Qy       373 LeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg---------SerAsp 389
Db      2373 AAGCAATCAGAGACTGTGGATCAAAATTCAGATTCTGATGAAATGCTAGCAATCCTCAAA 2432

Qy       390 GluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProVal 409
Db      2433 GAGGTGAGCAGGATGAGTCACAGTTCTTCTTCAGATACTGATATTAATGAAATTCATACA 2492

Qy       410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429
Db      2493 AACCATAAGACTTTGTATGATTTAAAGACTCAG---------GCGGGGAAAGAT------ 2537

Qy       430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp--------- 446
Db      2538 ------GATAAAGGAAAAAGGAAACGAAAAAGTTCTACATCTGGCTCAGATTTTGATACT 2591

Qy       447 ------GlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg----------- 460
Db      2592 AAAAAGGGCAAATCAGCTAAGAGCTCTATAATTTCTAAAAAGAAACGACAAACCCAGTCT 2651

Qy       461 -------------------------ThrGlySerValHisHisThrValAlaHisPro 471
Db      2652 GAGTCTTCTAATTATGACTCAGAATTAGAAAAAGAGATAAAGAGCATGAGTAAAATTGGT 2711

Qy       472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSer---ThrGlnHisAspAsp 490
Db      2712 GCTGCCAGAACCACCAAAAAAAGAATTCCAAATACAAAAGATTTTGACTCTTCTGAAGAT 2771

Qy       491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr---AspValCysGln 509
Db      2772 GAGAAACACAGCAAAAAAGGAATGGATAATCAAGGGCACAAAAATTTGAAGACCTCACAA 2831

Qy       510 HisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529
Db      2832 GAAGGATCATCTGATGATGCTGAAAGA---AAACAAGAGAGAGAGACTTTCTCTTCAGCA 2888

Qy       530 LysGly---------------------------------------------------Lys 532
Db      2889 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTAAGAGATCGACTTCCTAAGAAG 2948

Qy       533 ThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIle 552
Db      2949 CAGCAAGCAAGTGCTTCCACT------GATGGTGTCGATAAGCTTTCTGGGAAAGAGCAG 3002

Qy       553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHis 572
Db      3003 AGTTTTACTTCTTTGGAAGTTAGAAAAGTTGCTGAAACTAAAGAAAAG------AGCAAG 3056

Qy       573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592
Db      3057 CATCTCAAAACCAAAACATGTAAAAAAGTACAGGATGGCTTATCTGATATTGCAGAGAAA 3116

Qy       593 SerLeuProLysLysLys----------------------LysLysGlnLysLeuGlu 604
Db      3117 TTCCTAAAGAAAGACCAGAGCGATGAAACTTCTGAAGATGATAAAAAGCAGAGCAAAAAG 3176

Qy       605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db      3177 GGAACTGAAGAAAAAAAGAAA--------------CCTTCAGACTTTAAGAAAAAAGTA 3221

Qy       625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp--------- 641
```

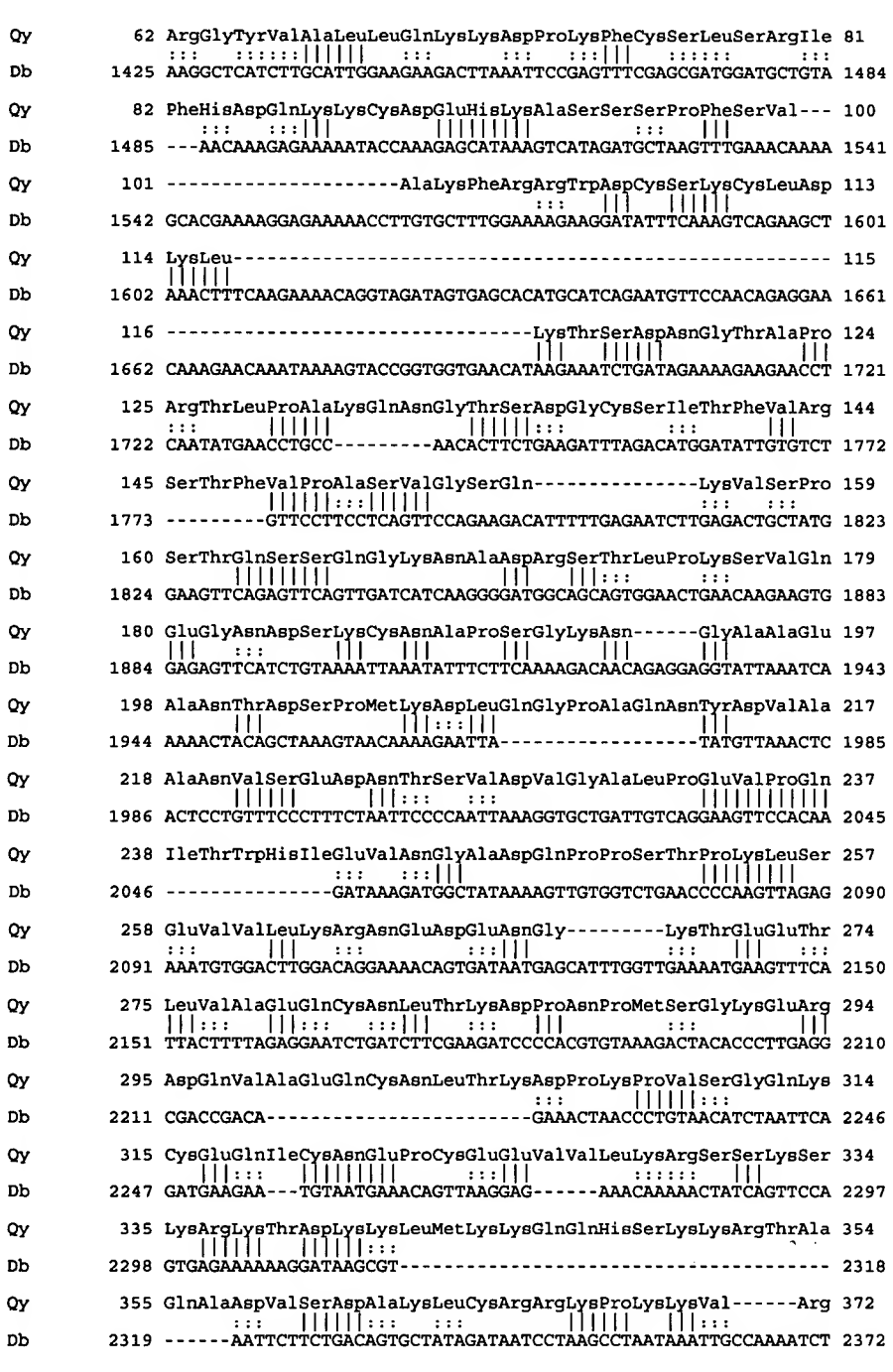
```
Db      3222 ATTAAA------ATGGAACAACAGTATGAATCTTCATCTGATGGCACTGAAAAGTTACCT 3275

Qy       642 --------------------------------IleAsnArgIleGlnSerLysThrThr 650
Db      3276 GAGCGAGAAGAAATTTGTCATTTTCCTAAGGGCATAAAACAAATTAAGAATGGAACAACT 3335

Qy       651 AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerServal 670
Db      3336 --------------------------------GATGGAGAAAAGAAAAGTAAAAAAATA 3362

Qy       671 PheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690
Db      3363 AGAGATAAAACTTCTAAAAAGAAGGATGAATTATCTGATTATGCTGAGAAGTCAA--CAG 3420

Qy       691 GlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGln 710
Db      3421 GGA---------------------------------------------------------- 3423

Qy       711 GluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeu 730
Db      3424 ----------------------AAGGAGATAGTTGTGACTCTT----------------- 3444

Qy       731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TrpGlyArgLy 750
Db      3445 ------------------CAGAGGATAAAAAGAGTAAGAATGGAGCATATGGTAGAGAGA 3486

Qy       750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765
Db      3487 AGAAAAGGTGCAAGTTGCTTGGAAAGAGTTCAAGGAAGAGACAAGA 3532
```

RESULT 14
US-10-934-998-25
; Sequence 25, Application US/10934998
; Publication No. US20050153917A1
; GENERAL INFORMATION:
; APPLICANT: AL-MAHMOOD, SALMAN
; APPLICANT: COLIN, SYLVIE
; APPLICANT: SCHNEIDER, CHRISTOPHE
; TITLE OF INVENTION: GENES INVOLVED IN REGULATING ANGIOGENESIS, PHARMACEUTICAL
; TITLE OF INVENTION: PREPARATIONS CONTAINING SAME AND APPLICATIONS THEREOF
; FILE REFERENCE: BMA-04-1206
; CURRENT APPLICATION NUMBER: US/10/934,998
; CURRENT FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: PCT/FR03/00695
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: FR02/02717
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: FR02/04546
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 10330
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: GS-N24
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GENBANK/U72937
; DATABASE ENTRY DATE: 2000-02-11
; RELEVANT RESIDUES: (1)..(10330)
US-10-934-998-25

Alignment Scores:

Pred. No.:	4.47e-08	Length:	10330
Score:	223.00	Matches:	172
Percent Similarity:	34.3%	Conservative:	108
Best Local Similarity:	21.1%	Mismatches:	312
Query Match:	4.0%	Indels:	225
DB:	9	Gaps:	33

US-09-828-068-2 (1-1057) x US-10-934-998-25 (1-10330)


Qy 62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81
Db 1425 AAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTA 1484
Qy 82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100
Db 1485 ---AACAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTGAAACAAAA 1541
Qy 101 --------------------AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113
Db 1542 GCACGAAAAGGAGAAAAACCTTGTGCTTTGGAAAAGAAGGATATTTCAAAGTCAGAAGCT 1601
Qy 114 LysLeu---------- 115
Db 1602 AAACTTTCAAGAAAACAGGTAGATAGTGAGCACATGCATCAGAATGTTCCAACAGAGGAA 1661
Qy 116 --------------------------------LysThrSerAspAsnGlyThrAlaPro 124
Db 1662 CAAAGAACAAATAAAAGTACCGGTGGTGAACATAAGAAATCTGATAGAAAAGAAGAACCT 1721
Qy 125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144
Db 1722 CAATATGAACCTGCC---------AACACTTCTGAAGATTTAGACATGGATATTGTGTCT 1772
Qy 145 SerThrPheValProAlaSerValGlySerGln--------------LysValSerPro 159
Db 1773 ---------GTTCCTTCCTCAGTTCCAGAAGACATTTTTGAGAATCTTGAGACTGCTATG 1823
Qy 160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179
Db 1824 GAAGTTCAGAGTTCAGTTGATCATCAAGGGGATGGCAGCAGTGGAACTGAACAAGAAGTG 1883
Qy 180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn------GlyAlaAlaGlu 197
Db 1884 GAGAGTTCATCTGTAAAATTAAATATTTCTTCAAAAGACAACAGAGGAGGTATTAAATCA 1943
Qy 198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217
Db 1944 AAAACTACAGCTAAAGTAACAAAAGAATTA-----------------TATGTTAAACTC 1985
Qy 218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237
Db 1986 ACTCCTGTTTCCCTTTCTAATTCCCCAATTAAAGGTGCTGATTGTCAGGAAGTTCCACAA 2045
Qy 238 IleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSer 257
Db 2046 --------------GATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAG 2090
Qy 258 GluValValLeuLysArgAsnGluAspGluAsnGly---------LysThrGluGluThr 274
Db 2091 AAATGTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCA 2150
Qy 275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294
Db 2151 TTACTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTTGAGG 2210
Qy 295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314
Db 2211 CGACCGACA----------GAAACTAACCCTGTAACATCTAATTCA 2246
Qy 315 CysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSer 334
Db 2247 GATGAAGAA---TGTAATGAAACAGTTAAGGAG------AAACAAAAACTATCAGTTCCA 2297
Qy 335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354
Db 2298 GTGAGAAAAAAGGATAAGCGT---------- 2318
Qy 355 GlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal------Arg 372
Db 2319 ------AATTCTTCTGACAGTGCTATAGATAATCCTAAGCCTAATAAATTGCCAAAATCT 2372
Qy 373 LeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg---------SerAsp 389
Db 2373 AAGCAATCAGAGACTGTGGATCAAAATTCAGATTCTGATGAAATGCTAGCAATCCTCAAA 2432
Qy 390 GluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProVal 409
Db 2433 GAGGTGAGCAGGATGAGTCACAGTTCTTCTTCAGATACTGATATTAATGAAATTCATACA 2492
Qy 410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429
Db 2493 AACCATAAGACTTTGTATGATTTAAAGACTCAG---------GCGGGGAAAGAT------ 2537
Qy 430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp--------- 446
Db 2538 ------GATAAAGGAAAAAGGAAACGAAAAAGTTCTACATCTGGCTCAGATTTTGATACT 2591
Qy 447 ------GlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg------------ 460
Db 2592 AAAAAGGGCAAATCAGCTAAGAGCTCTATAATTTCTAAAAAGAAACGACAAACCCAGTCT 2651
Qy 461 --------------------------ThrGlySerValHisHisThrValAlaHisPro 471
Db 2652 GAGTCTTCTAATTATGACTCAGAATTAGAAAAAGAGATAAAGAGCATGAGTAAAATTGGT 2711
Qy 472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSer---ThrGlnHisAspAsp 490
Db 2712 GCTGCCAGAACCACCAAAAAAAGAATTCCAAATACAAAAGATTTTGACTCTTCTGAAGAT 2771
Qy 491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr---AspValCysGln 509
Db 2772 GAGAAACACAGCAAAAAAGGAATGGATAATCAAGGGCACAAAAATTTGAAGACCTCACAA 2831
Qy 510 HisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529
Db 2832 GAAGGATCATCTGATGATGCTGAAAGA---AAACAAGAGAGAGAGACTTTCTCTTCAGCA 2888
Qy 530 LysGly---------- 532
Db 2889 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTAAGAGATCGACTTCCTAAGAAG 2948
Qy 533 ThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIle 552
Db 2949 CAGCAAGCAAGTGCTTCCACT------GATGGTGTCGATAAGCTTTCTGGGAAAGAGCAG 3002
Qy 553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHis 572
Db 3003 AGTTTTACTTCTTTGGAAGTTAGAAAAGTTGCTGAAACTAAAGAAAAG------AGCAAG 3056
Qy 573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592
Db 3057 CATCTCAAAACCAAAACATGTAAAAAAGTACAGGATGGCTTATCTGATATTGCAGAGAAA 3116
Qy 593 SerLeuProLysLysLys----------LysLysGlnLysLeuGlu 604
Db 3117 TTCCTAAAGAAAGACCAGAGCGATGAAACTTCTGAAGATGATAAAAAGCAGAGCAAAAAG 3176
Qy 605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624
Db 3177 GGAACTGAAGAAAAAAAGAAA----------CCTTCAGACTTTAAGAAAAAAGTA 3221
Qy 625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp--------- 641
Db 3222 ATTAAA------ATGGAACAACAGTATGAATCTTCATCTGATGGCACTGAAAAGTTACCT 3275
Qy 642 --------------------------------IleAsnArgIleGlnSerLysThrThr 650
Db 3276 GAGCGAGAAGAAATTTGTCATTTTCCTAAGGGCATAAAACAAATTAAGAATGGAACAACT 3335
Qy 651 AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVal 670
Db 3336 --------------------------------GATGGAGAAAAGAAAAGTAAAAAAATA 3362
Qy 671 PheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690

```
Db       3363 AGAGATAAAACTTCTAAAAAGAAGGATGAATTATCTGATTATGCTGAGAAGTCAA--CAG 3420

Qy        691 GlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGln 710

Db       3421 GGA------------------------------------------------------- 3423

Qy        711 GluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeu 730

Db       3424 --------------------AAGGAGATAGTTGTGACTCTT------------------ 3444

Qy        731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TrpGlyArgLy 750

Db       3445 ------------------CAGAGGATAAAAAGAGTAAGAATGGAGCATATGGTAGAGAGA 3486

Qy        750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765

Db       3487 AGAAAAGGTGCAAGTTGCTTGGAAAGAGTTCAAGGAAGAGACAAGA 3532

RESULT 15
US-09-968-007A-490
; Sequence 490, Application US/09968007A
; Publication No. US20040115625A1
; GENERAL INFORMATION:
;  APPLICANT: Ebner, Reinhard
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Usining Signa
;  TITLE OF INVENTION:  Gene Sets
;  FILE REFERENCE: 689290-71
;  CURRENT APPLICATION NUMBER: US/09/968,007A
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,172
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,173
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,278
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,294
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,295
;  PRIOR FILING DATE: 2000-10-02
;  PRIOR APPLICATION NUMBER: US/60/237,316
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 1001
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 490
;   LENGTH: 10452
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-968-007A-490

Alignment Scores:
Pred. No.:              4.55e-08      Length:          10452
Score:                  223.00        Matches:         172
Percent Similarity:     34.3%         Conservative:    108
Best Local Similarity:  21.1%         Mismatches:      312
Query Match:            4.0%          Indels:          225
DB:                     3             Gaps:            33

US-09-828-068-2 (1-1057) x US-09-968-007A-490 (1-10452)

Qy          62 ArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArgIle 81

Db        1547 AAGGCTCATCTTGCATTGGAAGAAGACTTAAATTCCGAGTTTCGAGCGATGGATGCTGTA 1606

Qy          82 PheHisAspGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerVal--- 100

Db        1607 ---AACAAAGAGAAAAATACCAAAGAGCATAAAGTCATAGATGCTAAGTTTGAAACAAAA 1663

Qy         101 --------------------AlaLysPheArgArgTrpAspCysSerLysCysLeuAsp 113

Db        1664 GCACGAAAAGGAGAAAAACCTTGTGCTTTGGAAAAGAAGGATATTTCAAAGTCAGAAGCT 1723

Qy         114 LysLeu---------------------------------------------------- 115

Db        1724 AAACTTTCAAGAAAACAGGTAGATAGTGAGCACATGCATCAGAATGTTCCAACAGAGGAA 1783

Qy         116 --------------------------------LysThrSerAspAsnGlyThrAlaPro 124

Db        1784 CAAAGAACAAATAAAAGTACCGGTGGTGAACATAAGAAATCTGATAGAAAAGAAGAACCT 1843

Qy         125 ArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArg 144

Db        1844 CAATATGAACCTGCC---------AACACTTCTGAAGATTTAGACATGGATATTGTGTCT 1894

Qy         145 SerThrPheValProAlaSerValGlySerGln---------------LysValSerPro 159

Db        1895 ---------GTTCCTTCCTCAGTTCCAGAAGACATTTTTGAGAATCTTGAGACTGCTATG 1945

Qy         160 SerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGln 179

Db        1946 GAAGTTCAGAGTTCAGTTGATCATCAAGGGGATGGCAGCAGTGGAACTGAACAAGAAGTG 2005

Qy         180 GluGlyAsnAspSerLysCysAsnAlaProSerGlyLysAsn------GlyAlaAlaGlu 197

Db        2006 GAGAGTTCATCTGTAAAATTAAATATTTCTTCAAAAGACAACAGAGGAGGTATTAAATCA 2065

Qy         198 AlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAla 217

Db        2066 AAAACTACAGCTAAAGTAACAAAAGAATTA-----------------TATGTTAAACTC 2107

Qy         218 AlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluValProGln 237

Db        2108 ACTCCTGTTTCCCTTTCTAATTCCCCAATTAAAGGTGCTGATTGTCAGGAAGTTCCACAA 2167

Qy         238 IleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSer 257

Db        2168 --------------GATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAG 2212

Qy         258 GluValValLeuLysArgAsnGluAspGluAsnGly---------LysThrGluGluThr 274

Db        2213 AAATGTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCA 2272

Qy         275 LeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg 294

Db        2273 TTACTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTTGAGG 2332

Qy         295 AspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLys 314

Db        2333 CGACCGACA-----------------------GAAACTAACCCTGTAACATCTAATTCA 2368

Qy         315 CysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSer 334

Db        2369 GATGAAGAA---TGTAATGAAACAGTTAAGGAG------AAACAAAAACTATCAGTTCCA 2419

Qy         335 LysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAla 354

Db        2420 GTGAGAAAAAAGGATAAGCGT------------------------------------- 2440

Qy         355 GlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal------Arg 372

Db        2441 ------AATTCTTCTGACAGTGCTATAGATAATCCTAAGCCTAATAAATTGCCAAAATCT 2494

Qy         373 LeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArg---------SerAsp 389

Db        2495 AAGCAATCAGAGACTGTGGATCAAAATTCAGATTCTGATGAAATGCTAGCAATCCTCAAA 2554

Qy         390 GluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIleProVal 409

Db        2555 GAGGTGAGCAGGATGAGTCACAGTTCTTCTTCAGATACTGATATTAATGAAATTCATACA 2614

Qy         410 ProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAspGlyLeu 429

Db        2615 AACCATAAGACTTTGTATGATTTAAAGACTCAG---------GCGGGGAAAGAT------ 2659

Qy         430 LysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAsp--------- 446
```

```
Db       2660 ------GATAAAGGAAAAAGGAAACGAAAAAGTTCTACATCTGGCTCAGATTTTGATACT 2713

Qy        447 ------GlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLysArg---------- 460

Db       2714 AAAAAGGGCAAATCAGCTAAGAGCTCTATAATTTCTAAAAAGAAACGACAAACCCAGTCT 2773

Qy        461 --------------------------ThrGlySerValHisHisThrValAlaHisPro 471

Db       2774 GAGTCTTCTAATTATGACTCAGAATTAGAAAAAGAGATAAAGAGCATGAGTAAAATTGGT 2833

Qy        472 AlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSer---ThrGlnHisAspAsp 490

Db       2834 GCTGCCAGAACCACCAAAAAAAGAATTCCAAATACAAAAGATTTTGACTCTTCTGAAGAT 2893

Qy        491 GluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThr---AspValCysGln 509

Db       2894 GAGAAACACAGCAAAAAAGGAATGGATAATCAAGGGCACAAAAATTTGAAGACCTCACAA 2953

Qy        510 HisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSer 529

Db       2954 GAAGGATCATCTGATGATGCTGAAAGA---AAACAAGAGAGAGAGACTTTCTCTTCAGCA 3010

Qy        530 LysGly-------------------------------------------------Lys 532

Db       3011 GAAGGCACAGTTGATAAAGACACGACCATCATGGAATTAAGAGATCGACTTCCTAAGAAG 3070

Qy        533 ThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIle 552

Db       3071 CAGCAAGCAAGTGCTTCCACT------GATGGTGTCGATAAGCTTTCTGGGAAAGAGCAG 3124

Qy        553 HisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHis 572

Db       3125 AGTTTTACTTCTTTGGAAGTTAGAAAAGTTGCTGAAACTAAAGAAAAG------AGCAAG 3178

Qy        573 SerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGln 592

Db       3179 CATCTCAAAACCAAAACATGTAAAAAAGTACAGGATGGCTTATCTGATATTGCAGAGAAA 3238

Qy        593 SerLeuProLysLysLys------------------------LysLysGlnLysLeuGlu 604

Db       3239 TTCCTAAAGAAAGACCAGAGCGATGAAACTTCTGAAGATGATAAAAAGCAGAGCAAAAAG 3298

Qy        605 ValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleValGluLeuLeu 624

Db       3299 GGAACTGAAGAAAAAAAGAAA--------------CCTTCAGACTTTAAGAAAAAAGTA 3343

Qy        625 AlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp--------- 641

Db       3344 ATTAAA------ATGGAACAACAGTATGAATCTTCATCTGATGGCACTGAAAAGTTACCT 3397

Qy        642 --------------------------------IleAsnArgIleGlnSerLysThrThr 650

Db       3398 GAGCGAGAAGAAATTTGTCATTTTCCTAAGGGCATAAAACAAATTAAGAATGGAACAACT 3457

Qy        651 AlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVal 670

Db       3458 --------------------------------GATGGAGAAAAGAAAAGTAAAAAAATA 3484

Qy        671 PheAspThrAsnSerGlnGlnLysSerLeuAlaSerGlnSerThrGlnLysGluLeuGln 690

Db       3485 AGAGATAAAACTTCTAAAAAGAAGGATGAATTATCTGATTATGCTGAGAAGTCAA--CAG 3542

Qy        691 GlyHisLeuAlaLeuThrThrGlnGluSerProHisProGlnAsnPheGlnSerThrGln 710

Db       3543 GGA---------------------------------------------------------- 3545

Qy        711 GluGlnGlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSerSerProLeu 730

Db       3546 ----------------------AAGGAGATAGTTGTGACTCTT----------------- 3566

Qy        731 PheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHis-TrpGlyArgLy 750

Db       3567 ------------------CAGAGGATAAAAAGAGTAAGAATGGAGCATATGGTAGAGAGA 3608

Qy        750 sAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArg 765

Db       3609 AGAAAAGGTGCAAGTTGCTTGGAAAGAGTTCAAGGAAGAGACAAGA 3654


Search completed: March  7, 2006, 14:42:12
Job time : 1816 secs
```

GenCore version 5.1.7

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 7, 2006, 14:08:44 ; Search time 620 Seconds
(without alignments)
3737.834 Million cell updates/sec

Title: US-09-828-068-2
Perfect score: 5526
Sequence: 1 MEIVAVDQEGARVVGTNCML.....NKNPADFTTISNDNEYMDYR 1057

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09828068/runat_06032006_163202_25316/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss03h
-USER=US09828068_@CGN_1_1_431_@runat_06032006_163202_25316 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	213	3.9	2290	9	US-11-072-175-21	Sequence 21, Appl
2	209.5	3.8	2920	12	US-11-124-368A-78	Sequence 78, Appl
3	206	3.7	9505	12	US-11-120-925-2	Sequence 2, Appli
4	202.5	3.7	4174	12	US-11-124-368A-75	Sequence 75, Appl
5	202	3.7	5545	7	US-10-330-773-84	Sequence 84, Appl
6	202	3.7	6189	8	US-10-909-125-825	Sequence 825, App
7	200.5	3.6	3989	12	US-11-124-368A-70	Sequence 70, Appl
8	200.5	3.6	4143	12	US-11-124-368A-79	Sequence 79, Appl
9	199	3.6	2782	12	US-11-124-368A-81	Sequence 81, Appl
10	198.5	3.6	3076	12	US-11-124-368A-71	Sequence 71, Appl
11	197	3.6	2943	12	US-11-124-368A-69	Sequence 69, Appl
12	196	3.5	3122	12	US-11-124-368A-76	Sequence 76, Appl
13	195	3.5	2004	8	US-10-821-234-625	Sequence 625, App
14	195	3.5	4075	12	US-11-124-368A-80	Sequence 80, Appl
15	189	3.4	2791	12	US-11-124-368A-74	Sequence 74, Appl
16	189	3.4	6266	8	US-10-955-054A-24	Sequence 24, Appl
17	189	3.4	6270	12	US-11-124-367A-185	Sequence 185, App
18	188.5	3.4	4834	12	US-11-000-688-1005	Sequence 1005, Ap
19	187	3.4	2998	12	US-11-124-368A-73	Sequence 73, Appl
20	187	3.4	3151	12	US-11-124-368A-72	Sequence 72, Appl
21	183.5	3.3	26241	12	US-11-098-686-8887	Sequence 8887, Ap
22	183.5	3.3	194553	12	US-11-098-686-8738	Sequence 8738, Ap
23	182.5	3.3	3615	8	US-10-793-626-3565	Sequence 3565, Ap
24	182.5	3.3	8503	12	US-11-124-368A-45	Sequence 45, Appl
25	182	3.3	7748	12	US-11-136-527-3397	Sequence 3397, Ap
26	181.5	3.3	3344	8	US-10-793-626-3504	Sequence 3504, Ap
27	181.5	3.3	3840	8	US-10-793-626-3187	Sequence 3187, Ap
28	181.5	3.3	4042	8	US-10-793-626-3701	Sequence 3701, Ap
29	181.5	3.3	10300	8	US-10-947-249-122	Sequence 122, App
30	180.5	3.3	3257	12	US-11-124-368A-77	Sequence 77, Appl
31	180.5	3.3	4023	12	US-11-098-686-9790	Sequence 9790, Ap
c 32	180.5	3.3	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
33	180	3.3	8512	12	US-11-124-368A-46	Sequence 46, Appl
34	179.5	3.2	3161	12	US-11-124-368A-68	Sequence 68, Appl
35	178.5	3.2	134499	12	US-11-117-187-192	Sequence 192, App
36	176.5	3.2	15071	8	US-10-793-626-2963	Sequence 2963, Ap
37	175.5	3.2	8831	12	US-11-173-791-10	Sequence 10, Appl
c 38	175.5	3.2	116856	12	US-11-143-980-1	Sequence 1, Appli
c 39	174.5	3.2	3598	8	US-10-793-626-4216	Sequence 4216, Ap
c 40	174	3.1	3543	8	US-10-793-626-4149	Sequence 4149, Ap
41	174	3.1	3867	8	US-10-793-626-3769	Sequence 3769, Ap
42	173.5	3.1	3255	7	US-10-330-773-91	Sequence 91, Appl
43	173.5	3.1	3285	8	US-10-793-626-3153	Sequence 3153, Ap
44	173.5	3.1	3730	8	US-10-793-626-3499	Sequence 3499, Ap
45	171.5	3.1	2366	9	US-11-072-512-912	Sequence 912, App

ALIGNMENTS

RESULT 1
US-11-072-175-21
; Sequence 21, Application US/11072175
; Publication No. US20060029944A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THA
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273A CIP
; CURRENT APPLICATION NUMBER: US/11/072,175
; CURRENT FILING DATE: 2005-03-05
; PRIOR APPLICATION NUMBER: US 60/406,385
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 10/648,593
; PRIOR FILING DATE: 2003-08-26
; NUMBER OF SEQ ID NOS: 571
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 2290
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-175-21

Alignment Scores:
Pred. No.: 1.1e-07 Length: 2290
Score: 213.00 Matches: 168
Percent Similarity: 35.9% Conservative: 103

Best Local Similarity: 22.3% Mismatches: 292
Query Match: 3.9% Indels: 192
DB: 9 Gaps: 35

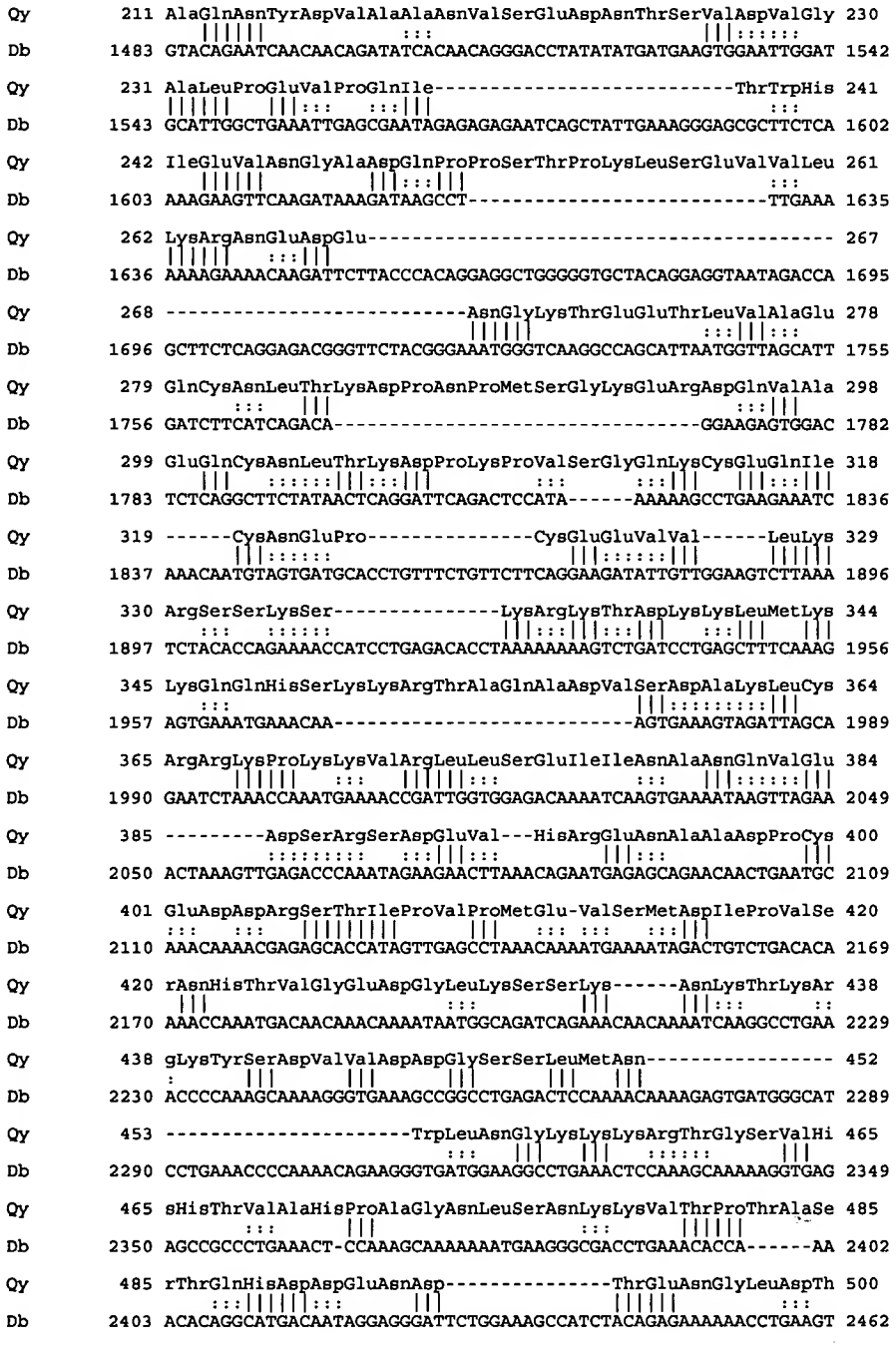
US-09-828-068-2 (1-1057) x US-11-072-175-21 (1-2290)

```
Qy          58 HisPheSerIleArgGlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSer 77
Db          60 CACTTTAAA---------TACTTTTATCTACTGAATCGCAAGAAAATTGGGTTTCTGGTG 110

Qy          78 LeuSerArgIlePheHisAspGlnLys------------LysCysAspGluHis------ 91
Db         111 CTGTCT---GTGGTTCATGAGAAAAAATCCCAAGAAGGAAAGCCAAAAGAACACACAGAG 167

Qy          92 --------------LysAlaSerSerSerProPheSerValAlaLysPheArgArgTrp 106
Db         168 CCAAAAAGCCTACCCAAGCAGGCATCAGATACAGGAAGTAACGATGCTCACAATAAAAAA 227

Qy         107 AspCysSerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThr 126
Db         228 GCAGTTTCCAGATCAGCTGAACAG---CAGCCATCAGAGAAATCAACAGAACCAAAGACT 284

Qy         127 LeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThr 146
Db         285 AAACCACAAGACATG---------------------------------------------- 299

Qy         147 PheValProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGly 166
Db         300 ---------ATTTCTGCTGGTGGAGAGAGTGTTGCTGGTATCACTGCAATATCT---GGC 347

Qy         167 LysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGly------AsnAspSer 184
Db         348 AAGCCGGGTGACAAGAAAAAAGAAAAGAAATCATTAACCCCAGCTGTGCCAGTTGAATCT 407

Qy         185 LysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMet 204
Db         408 AAACCGGATAAACCATCGGGAAAGTCAGGC---ATGGATGCTGCTTTGGATGACTTAATA 464

Qy         205 LysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsn 224
Db         465 GATACTTTAGGAGGACCTGAAGAA--------------------ACTGAAGAAGAAAAT 503

Qy         225 ThrSerValAspValGlyAlaLeuProGluVal------ProGlnIleThrTrpHisIle 242
Db         504 ACAACGTAT---------ACTGGACCAGAAGTTTCAGATCCAATGAGTTCCACCTACATA 554

Qy         243 GluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeuLys 262
Db         555 GAGGAATTGGGTAAAAGAGAAGTCACAATTCCTCCAAAATATAGGGAACTATTGGCTAAA 614

Qy         262 ---------------------------------------------------------- 262
Db         615 CCCATAGGGCCAGATGATGCTATAGACGCCTTGTCATCTGACTTCACCTGTGGGTCGCCT 674

Qy         263 ---------ArgAsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGln 279
Db         675 ACAGCTGCTGGAAAGAAAACTGAAAAAGAGGAATCTACAGAAGTTTTAAAAGCTCAGTCA 734

Qy         280 CysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg--------------- 294
Db         735 GCAGGGACAGTCAGAAGTGCTGCTCCACCCCAAGAGAAGAAAAGAAAGGTGGAGAAGGAT 794

Qy         295 ---------AspGlnValAlaGluGln-CysAsnLeuThrLysAspProLysProValSe 311
Db         795 ACAATGAGTGATCAAGCACTCGAGGCTCTGTCGGCTTCACTGGGCACCC----------- 843

Qy         311 rGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSe 331
Db         844 -GGCAAGCAG--------------AACCTGAGCTCGACCTCCGCTCAATTAAGGAAGTC 887

Qy         331 rSerLysSerLysArgLysThrAspLysLys-----------------LeuMetLysLy 345
Db         888 GATGAGGCAAAAGCTAAAG------AAGAAAAACTAGAGAAGTGTGGTGAGGATGATGAA 941

Qy         345 sGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysAr 365
Db         942 ACAATCCCATCT-GAGTACAGATTAAAACCAGCCACGGATAAAGATGGAAAACCACTATT 1000

Qy         365 gArgLysProLys---------LysValArgLeuLeuSerGluIleIleAsnAlaAsnGl 382
Db        1001 GCCAGAGCCTGAAGAAAAACCCAAGCCTCGGAGTGAATCAGAACTCATTGATGAACTTTC 1060

Qy         382 nValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro-------- 399
Db        1061 AGAAGATTTTGACCGGTCTGAA------TGTAAAGAGAAACCATCTAAGCCAACTGAAAA 1114

Qy         400 -CysGluAspAspArgSerThrIleProValProMetGluValSerMet----------- 415
Db        1115 GACAGAAGAATCTAAGGCCGCTGCTCCAGCTCCTGTGTCGGAGGCTGTGTGTCGGACCTC 1174

Qy         416 ----------------------------AspIleProValSerAsnHisThrValGlyGl 426
Db        1175 CATGTGTAGTATACAGTCAGCACCCCCTGAGCCGGCTACCTTGAAGGGCACAGTGCCAGA 1234

Qy         426 uAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAs 446
Db        1235 TGATGCTGTAGAAGCCTTGGCTGATAGCCTGGGGAAAAAGGAAGCAGAT---CCAGAAGA 1291

Qy         446 pGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLys--------------ArgTh 461
Db        1292 TGGAAAACCTGTGATGGATAAAGTCAAGGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCT 1351

Qy         461 rGlySerValHisHisThrVal--------------AlaHisProAlaGlyAsnLeuSe 476
Db        1352 TGGTGAAAAAGAAGAAACAATTCCTCCTGATTATAGATTAGAAGAGGTCAAGGATAAAGA 1411

Qy         476 rAsnLysLysValThrProThrAlaSerThrGlnHis---------------------- 488
Db        1412 TGGAAAGCCACTCCTGCCAAAAGAGTCTAAGGAACAGCTTCCACCCATGAGTGAAGACTT 1471

Qy         489 -------------------AspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMe 502
Db        1472 CCTTCTGGATGCTTTGTCTGAGGACTTCTCTGGTCCACAAAATGCTTCATCTCTTAAATT 1531

Qy         502 tHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLy 522
Db        1532 TGAAGATGCTAAACTTGCTGCTGCCATCTCTGAAGTGGTTTCCCAA-------------- 1577

Qy         522 sGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGl 542
Db        1578 ------------------------------------ACCCCAGCTTCAACGACCCAAGC 1600

Qy         542 yGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGl 562
Db        1601 TGGAGCCCCACCCCGTGAT--------------------ACCTCGCAGAGTGACAAAGA 1639

Qy         562 nMetGluThrGluAsnSerValLeuSerHisSer---AlaLysValSerProAlaGluHi 581
Db        1640 CCTCGATGATGCCTTGGATAAACTCTCTGACAGTCTAGGACAAAGGCAGCCTGACCCAGA 1699

Qy         581 sAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGl 601
Db        1700 TGAGAACAAACCAATGGAAGAT--------------------AAAGTAAAGGAAAAAGC 1738

Qy         601 nLysLeuGluValThrArgGluLysGlnThrMetIleAspAsp---IleProMetAspIl 620
Db        1739 TAAAGCTGAA---CATAGAGACAAGCTTGGAGAAAGAGATGACACTATCCCACCTGAATA 1795

Qy         620 eValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSe 640
Db        1796 CAGACATCTCCTGGATGATAATGGACAGGACAAACCAGTGAAGCCACCTACAAAGAAATC 1855

Qy         640 rAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAl 660
Db        1856 AGAG------------GATTCAAAGAAACCTGCAGATGACCAAGACCCCATTGATGCTCT 1903
```

```
Qy        660 aLysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln-------- 677
Db       1904 CTCAGGAGATCTGGACAGCTGTCCCTCCACTACAGAAACCTCACAGAACACAGCAAAGGA 1963

Qy        678 -------LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeu----- 693
Db       1964 TAAGTGCAAGAAGGCTGCTTCCAGCTCCAAAGCACCTAAGAATGGAGGTAAAGCGAAGGA 2023

Qy        694 ----AlaLeuThrThrGlnGluSerProHisProGlnAsn 705
Db       2024 TTCAGCAAAGACAACAGAGGAAACTTCCAAGCCAAAAGAT 2063

RESULT 2
US-11-124-368A-78
; Sequence 78, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78
;   LENGTH: 2920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-78

Alignment Scores:
Pred. No.:              2.82e-07      Length:          2920
Score:                  209.50        Matches:         182
Percent Similarity:     35.7%         Conservative:    114
Best Local Similarity:  22.0%         Mismatches:      320
Query Match:            3.8%          Indels:          214
DB:                     12            Gaps:            37

US-09-828-068-2 (1-1057) x US-11-124-368A-78 (1-2920)

Qy         24 GlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGlnAsp------ 41
Db         62 GGGACCAGAGCAGTGCTGGGGCCGGGGCGGGTCACCGGGACTCCCCGCCAGGCCTCCCCG 121

Qy         42 ------AlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPhe 59
Db        122 CCACTCTCCGCGGCGCATTCCGGGAGGCAGCGGCCGCAGCGGCCTCGC------CATGTC 175

Qy         60 SerIleArgGlyTyrVal-AlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSe 79
Db        176 CCAGCCCGGCCAGAAGCCCGCCGCCTCCCCGCGG------CCCCGGCGAGCAGCCGCCGC 229

Qy         79 rArgIlePheHis-----------------------------------------------As 84
Db        230 CCGCCGCACCCATGAGCATGTCAGTGAAAAAACCAGTGAATCGCCTTCCAAACCAGGAGA 289

Qy         84 pGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSerValAlaLysPheAr 104
Db        290 AAAGAAAGGATCAGATGAGGTGTCAGCTTCCTCTGGTGCAACCAGCAAGTCTTCCAGTAT 349

Qy        104 gArgTrpAspCysSerLysCysLeu---AspLysLeuLysThrSerAspAsnGlyThrAl 123
Db        350 GAATCCCACAGAAACCAAGGCTGTAAAAACAGAACCTGAGAAGAAGTCACAGTCAACCAA 409

Qy        123 aProArgThrLeuPro------AlaLysGlnAsnGlyThrSerAspGlyCysSerIleTh 141
Db        410 GCCAAAAAGCCTACCCAAGCAGGCATCAGATACAGGAAGTAACGATGCTCACAATAAAAA 469

Qy        141 rPheVal-----------------------------ArgSerThrPheValProAla-- 150
Db        470 AGCAGTTTCCAGATCAGCTGAACAGCAGCCATCAGAGAAATCAACAGAACCAAAGACTAA 529

Qy        151 ----------------SerValGlySerGlnLysValSerProSerThrGlnSerSerGl 165
Db        530 ACCACAAGACATGATTTCTGCTGGTGGAGAGAGTGTTGCTGGTATCACTGCAATATCT-- 587

Qy        165 nGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGly------AsnAs 183
Db        588 -GGCAAGCCGGGTGACAAGAAAAAAGAAAAGAAATCATTAACCCCAGCTGTGCCAGTTGA 646

Qy        183 pSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerPr 203
Db        647 ATCTAAACCGGATAAACCATCGGGAAAGTCAGGC---ATGGATGCTGCTTTGGATGACTT 703

Qy        203 oMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAs 223
Db        704 AATAGATACTTTAGGAGGACCTGAAGAA--------------------ACTGAAGAAGA 742

Qy        223 pAsnThrSerValAspValGlyAlaLeuProGluVal------ProGlnIleThrTrpHi 241
Db        743 AAATACAACGTAT---------ACTGGACCAGAAGTTTCAGATCCAATGAGTTCCACCTA 793

Qy        241 sIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLe 261
Db        794 CATAGAGGAATTGGGTAAAAGAGAAGTCACAATTCCTCCAAAATATAGGGAACTATTGGC 853

Qy        261 uLysArg---------------------------------------------------- 263
Db        854 TAAAAAGGAAGGGATCACAGGGCCTCCTGCAGACTCTTCGAAACCCATAGGGCCAGATGA 913

Qy        264 ----------------------------------------------------AsnGl 265
Db        914 TGCTATAGACGCCTTGTCATCTGACTTCACCTGTGGGTCGCCTACAGCTGCTGGAAAGAA 973

Qy        265 uAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAs 285
Db        974 AACTGAAAAAGAGGAATCTACAGAAGTTTTAAAAGCTCAGTCAGCAGGGACAGTCAGAAG 1033

Qy        285 pProAsnProMetSerGlyLysGluArg-----------------------AspGlnVa 297
Db       1034 TGCTGCTCCACCCCAAGAGAAGAAAAGAAAGGTGGAGAAGGATACAATGAGTGATCAAGC 1093

Qy        297 lAlaGluGln-CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluG 317
Db       1094 ACTCGAGGCTCTGTCGGCTTCACTGGGCACCC-----------GGCAAGCAG------- 1134

Qy        317 lnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSerLysArgL 337
Db       1135 --------AACCTGAGCTCGACCTCCGCTCAATTAAGGAAGTCGATGAGGCAAAAGCTAA 1186

Qy        337 ysThrAspLysLys-----------------LeuMetLysLysGlnGlnHisSerLysL 351
Db       1187 AG------AAGAAAAACTAGAGAAGTGTGGTGAGGATGATGAAACAATCCCATCT-GAGT 1239

Qy        351 ysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLys---- 369
Db       1240 ACAGATTAAAACCAGCCACGGATAAAGATGGAAAACCACTATTGCCAGAGCCTGAAGAAA 1299

Qy        370 -----LysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgS 388
Db       1300 AACCCAAGCCTCGGAGTGAATCAGAACTCATTGATGAACTTTCAGAAGATTTTGACCGGT 1359

Qy        388 erAspGluValHisArgGluAsnAlaAlaAspPro---------CysGluAspAspArgS 405
Db       1360 CTGAA------TGTAAAGAGAAACCATCTAAGCCAACTGAAAAGACAGAAGAATCTAAGG 1413

Qy        405 erThrIleProValProMetGluValSerMet--------------------------- 415
Db       1414 CCGCTGCTCCAGCTCCTGTGTCGGAGGCTGTGTGTCGGACCTCCATGTGTAGTATACAGT 1473
```

```
Qy        416 -----------AspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerS 432
Db       1474 CAGCACCCCCTGAGCCGGCTACCTTGAAGGGCACAGTGCCAGATGATGCTGTAGAAGCCT 1533

Qy        432 erLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetA 452
Db       1534 TGGCTGATAGCCTGGGGAAAAAGGAAGCAGAT---CCAGAAGATGGAAAACCTGTGATGG 1590

Qy        452 snTrpLeuAsnGlyLysLysLys--------------ArgThrGlySerValHisHisT 467
Db       1591 ATAAAGTCAAGGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCTTGGTGAAAAAGAAGAAA 1650

Qy        467 hrVal--------------AlaHisProAlaGlyAsnLeuSerAsnLysLysValThrP 482
Db       1651 CAATTCCTCCTGATTATAGATTAGAAGAGGTCAAGGATAAAGATGGAAAGCCACTCCTGC 1710

Qy        482 roThrAlaSerThrGlnHis---------------------------------------- 488
Db       1711 CAAAAGAGTCTAAGGAACAGCTTCCACCCATGAGTGAAGACTTCCTTCTGGATGCTTTGT 1770

Qy        489 --AspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValC 508
Db       1771 CTGAGGACTTCTCTGGTCCACAAAATGCTTCATCTCTTAAATTTGAAGATGCTAAACTTG 1830

Qy        508 ysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyL 528
Db       1831 CTGCTGCCATCTCTGAAGTGGTTTCCCAA------------------------------ 1859

Qy        528 euSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgA 548
Db       1860 --------------------ACCCCAGCTTCAACGACCCAAGCTGGAGCCCCACCCCGTG 1899

Qy        548 snGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnS 568
Db       1900 AT--------------------ACCTCGCAGAGTGACAAAGACCTCGATGATGCCTTGG 1938

Qy        568 erValLeuSerHisSer---AlaLysValSerProAlaGluHisAspIleGlnIleMetS 587
Db       1939 ATAAACTCTCTGACAGTCTAGGACAAAGGCAGCCTGACCCAGATGAGAACAAACCAATGG 1998

Qy        587 erAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrA 607
Db       1999 AAGAT--------------------AAAGTAAAGGAAAAAGCTAAAGCTGAA---CATA 2034

Qy        607 rgGluLysGlnThrMetIleAspAsp---IleProMetAspIleValGluLeuLeuAlaL 626
Db       2035 GAGACAAGCTTGGAGAAAGAGATGACACTATCCCACCTGAATACAGACATCTCCTGGATG 2094

Qy        626 ysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleG 646
Db       2095 ATAATGGACAGGACAAACCAGTGAAGCCACCTACAAAGAAATCAGAG-----------G 2142

Qy        646 lnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspT 666
Db       2143 ATTCAAAGAAACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCAGGAGATCTGGACA 2202

Qy        666 yrAlaSerSerValPheAspThrAsnSerGlnGln--------------LysSerLeuA 681
Db       2203 GCTGTCCCTCCACTACAGAAACCTCACAGAACACAGCAAAGGATAAGTGCAAGAAGGCTG 2262

Qy        681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeu---------AlaLeuThrThrG 698
Db       2263 CTTCCAGCTCCAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAGCAAAGACAACAG 2322

Qy        698 lnGluSerProHisProGlnAsn 705
Db       2323 AGGAAACTTCCAAGCCAAAAGAT 2345

RESULT 3
US-11-120-925-2
; Sequence 2, Application US/11120925
; Publication No. US20060003354A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Krantz, Ian D.
;  APPLICANT:  Jackson, Laird G.
;  TITLE OF INVENTION: Methods and Compositions for the
;  TITLE OF INVENTION:  Diagnosis of Cornelia De Lange Syndrome
;  FILE REFERENCE: 3460-CHOP.0235US
;  CURRENT APPLICATION NUMBER: US/11/120,925
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: US 60/567,756
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 9505
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-11-120-925-2

Alignment Scores:
Pred. No.:                2.07e-06      Length:          9505
Score:                    206.00        Matches:         232
Percent Similarity:       33.2%         Conservative:    165
Best Local Similarity:    19.4%         Mismatches:      415
Query Match:              3.7%          Indels:          385
DB:                       12            Gaps:            56

US-09-828-068-2 (1-1057) x US-11-120-925-2 (1-9505)

Qy         29 AlaProValLeuGluLeuThrAlaThrProArg--------------------------- 39
Db        962 GCTCCCCTGCTGGAAGTGAA-GGAACTCCTAAAGGCTCAAGACCACCTTTAATCCTACAA 1020

Qy         40 --------------------------GlnAspAlaAlaAlaGluAlaGlyValAspGlu 50
Db       1021 TCTCAGTCTCTACCTTGTTCATCACCTCGAGATGTTCCACCAGATATCTTGCTAGATTCT 1080

Qy         51 ProAlaGlnHisGlnCysGluHisPheSerIleArgGlyTyrValAlaLeuLeuGlnLys 70
Db       1081 CCAGAAAGAAAA-------------------------------------------CAAAAG 1098

Qy         71 LysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLysCysAspGlu 90
Db       1099 AAGCAGAAGAAAATG-----------AAATTAGGCAAGGATGAAAAAGAGCAGAGTGAG 1146

Qy         91 HisLysAlaSer------------SerSerProPheSerValAlaLysPheArgArgTrp 106
Db       1147 AAAGCGGCAATGTATGATATAATTAGTTCTCCATCCAAG--------------------- 1185

Qy        107 AspCysSerLysCys---------LeuAspLysLeuLysThrSerAspAsnGlyThrAla 123
Db       1186 GACTCTACTAAACTTACATTAAGACTTTCTCGTGTAAGGTCTTCAGACATGGAC------ 1239

Qy        124 ProArgThrLeuProAlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheVal 143
Db       1240 -----------------CAGCAAGAGGATATGATTTCTGGTGTGGAAAATAGCAATGTT 1281

Qy        144 ArgSerThrPheValProAlaSerVal---------GlySerGlnLysValSerProSer 160
Db       1282 TCAGAAAATGATATTCCTTTTAATGTGCAGTACCCAGGACAGACTTCAAAAACACCCATT 1341

Qy        161 ThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGlu 180
Db       1342 ACT-----------------------------------------CCACAAGATATAAACCGC 1362

Qy        181 Gly---AsnAspSerLysCysAsnAla-----------------------ProSerGly 191
Db       1363 CCACTAAATGCTGCTCAATGTTTGTCGCAGCAAGAACAAACAGCATTCCTTCCAGCAAAT 1422

Qy        192 LysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGln---GlyPro 210
Db       1423 CAAGTGCCTGTTTTACAACAGAACACTTCAGTTGCTGCAAAACAACCCCAGACTTCTGTG 1482
```


Qy 211 AlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGly 230
Db 1483 GTACAGAATCAACAACAGATATCACAACAGGGACCTATATATGATGAAGTGGAATTGGAT 1542
Qy 231 AlaLeuProGluValProGlnIle----------ThrTrpHis 241
Db 1543 GCATTGGCTGAAATTGAGCGAATAGAGAGAGAATCAGCTATTGAAAGGGAGCGCTTCTCA 1602
Qy 242 IleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeu 261
Db 1603 AAAGAAGTTCAAGATAAAGATAAGCCT----------TTGAAA 1635
Qy 262 LysArgAsnGluAspGlu---------- 267
Db 1636 AAAAGAAAACAAGATTCTTACCCACAGGAGGCTGGGGGTGCTACAGGAGGTAATAGACCA 1695
Qy 268 ----------AsnGlyLysThrGluGluThrLeuValAlaGlu 278
Db 1696 GCTTCTCAGGAGACGGGTTCTACGGGAAATGGGTCAAGGCCAGCATTAATGGTTAGCATT 1755
Qy 279 GlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArgAspGlnValAla 298
Db 1756 GATCTTCATCAGACA----------GGAAGAGTGGAC 1782
Qy 299 GluGlnCysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIle 318
Db 1783 TCTCAGGCTTCTATAACTCAGGATTCAGACTCCATA------AAAAAGCCTGAAGAAATC 1836
Qy 319 ------CysAsnGluPro----------CysGluGluValVal------LeuLys 329
Db 1837 AAACAATGTAGTGATGCACCTGTTTCTGTTCTTCAGGAAGATATTGTTGGAAGTCTTAAA 1896
Qy 330 ArgSerSerLysSer----------LysArgLysThrAspLysLysLeuMetLys 344
Db 1897 TCTACACCAGAAAACCATCCTGAGACACCTAAAAAAAAGTCTGATCCTGAGCTTTCAAAG 1956
Qy 345 LysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCys 364
Db 1957 AGTGAAATGAAACAA----------AGTGAAAGTAGATTAGCA 1989
Qy 365 ArgArgLysProLysLysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGlu 384
Db 1990 GAATCTAAACCAAATGAAAACCGATTGGTGGAGACAAAATCAAGTGAAAATAAGTTAGAA 2049
Qy 385 ---------AspSerArgSerAspGluVal---HisArgGluAsnAlaAlaAspProCys 400
Db 2050 ACTAAAGTTGAGACCCAAATAGAAGAACTTAAACAGAATGAGAGCAGAACAACTGAATGC 2109
Qy 401 GluAspAspArgSerThrIleProValProMetGlu-ValSerMetAspIleProValSe 420
Db 2110 AAACAAAACGAGAGCACCATAGTTGAGCCTAAACAAAATGAAAATAGACTGTCTGACACA 2169
Qy 420 rAsnHisThrValGlyGluAspGlyLeuLysSerSerLys------AsnLysThrLysAr 438
Db 2170 AAACCAAATGACAACAAACAAAATAATGGCAGATCAGAAACAACAAAATCAAGGCCTGAA 2229
Qy 438 gLysTyrSerAspValValAspAspGlySerSerLeuMetAsn---------- 452
Db 2230 ACCCCAAAGCAAAAGGGTGAAAGCCGGCCTGAGACTCCAAAACAAAAGAGTGATGGGCAT 2289
Qy 453 ----------TrpLeuAsnGlyLysLysLysArgThrGlySerValHi 465
Db 2290 CCTGAAACCCCAAAACAGAAGGGTGATGGAAGGCCTGAAACTCCAAAGCAAAAAGGTGAG 2349
Qy 465 sHisThrValAlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSe 485
Db 2350 AGCCGCCCTGAAACT-CCAAAGCAAAAAAATGAAGGGCGACCTGAAACACCA------AA 2402
Qy 485 rThrGlnHisAspAspGluAsnAsp----------ThrGluAsnGlyLeuAspTh 500
Db 2403 ACACAGGCATGACAATAGGAGGGATTCTGGAAAGCCATCTACAGAGAAAAAACCTGAAGT 2462
Qy 500 rAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSe 520
Db 2463 GTCTAAACATAAACAAGATACTAAATCTGACTCACCTCGGTTAAAATCAGAACGAGCTGA 2522
Qy 520 rSer----------LysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSe 535
Db 2523 AGCCTTAAAGCAGAGACCTGATGGGCGATCTGTTTCTGAGTCACTAAGACGTGACCATGA 2582
Qy 535 r----------AlaAlaSerThrLysTyrGlyGlyGluSerThrAr 547
Db 2583 TAATAAACAAAAATCAGATGACAGGGGTGAATCAGAGCGACATCGAGGGGATCAGTCTAG 2642
Qy 547 g----------AsnGlyGlnAsnIleHisValLeuSerAlaGl 558
Db 2643 GGTTCGAAGACCAGAAACATTGAGATCCTCTAGTAGAAATGAACATGGCATTAAATCTGA 2702
Qy 558 u---------- 558
Db 2703 TAGTTCAAAAACTGATAAACTAGAACGAAAACACAGGCACGAATCAGGGGACTCAAGGGA 2762
Qy 559 ----------AspGlnCysGlnMetGluThrGluAs 567
Db 2763 AAGACCATCTTCTGGGGAACAAAAATCAAGACCTGACAGTCCTCGTGTTAAACAAGGAGA 2822
Qy 567 nSerValLeuSerHisSerAlaLysVal---------SerProAlaGluHisAspIleGl 584
Db 2823 TTCTAATAAATCAAGATCTGATAAACTTGGTTTTAAATCACCAACTAGTAAAGATGACAA 2882
Qy 584 nIleMetSerAsp---------- 588
Db 2883 AAGGACAGAGGGTAACAAGAGTAAAGTAGACACTAATAAAGCACACCCTGACAATAAGGC 2942
Qy 589 ----------LeuHisGluGlnSerLeuPr 595
Db 2943 AGAATTTCCAAGTTATTTGTTGGGGGGCAGGTCTGGTGCGTTGAAAAATTTTGTCATTCC 3002
Qy 595 oLysLysLysLysGlnLysLeuGluValThrArgGluLysGlnThrMet---IleAs 614
Db 3003 GAAAATCAAGAGGGATAAAGATGGCAATGTTACTCAGGAGACAAAGAAAATGGAAATGAA 3062
Qy 614 pAspIleProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMe 634
Db 3063 AGGAGAGCCGAAAGACAAAGTAGAAAAAATAGGATTA---------- 3099
Qy 634 tThrGluThrAspCysSerAspIleAsnArg----------IleGlnSerLy 648
Db 3100 ----------GTTGAAGATCTAAATAAAGGAGCTAAGCCTGTAGTTGTGCTACAAAA 3146
Qy 648 sThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSe 668
Db 3147 ACTGTCTTTGGATGAT------GTTCAGAAACTTATTAAAGATAGAGAGGACAAATCGAG 3200
Qy 668 rSerValPhe----------AspThrAsnSerGlnGlnLysSerLeuAlaSerGl 683
Db 3201 AAGTTCCCTTAAACCTATCAAGAATAAACCATCAAAGTCAAATAAAGGTAGTATAGATCA 3260
Qy 683 nSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSerProHisPr 703
Db 3261 ATCAGTGTTAAAAGAATTACCCCCTGAACTC---CTGGCAGAAATTGAGTCCACCATGCC 3317
Qy 703 oGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGluGluMetValTh 723
Db 3318 ACTTTGTGAACGTGTGAAAATGAACAAACGCAAGCGTAGCACAGTTAATGAAAAGCCAAA 3377
Qy 723 rIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaPr 743
Db 3378 ATATGCTGAAATCAGTTCAGATGAAGATAATGATAGTGATGAAGCTTTTGAATCCTCTAG 3437
Qy 743 oThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaTh 763
Db 3438 GAAACGACAT------AAAAAAGATGATGATAAA---GCTTGGGAATATGAAGAGCGTGA 3488
Qy 763 rThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaVa 783

```
Db       3489 CAGAAGAAGCTCT--------------GGGGATCATAGGAGAAGTGGC----------- 3522

Qy        783 lAspLeuThrSerThrHisValMetGlySerSerSer---------AsnTyrAlaSerAr 800

Db       3523 ----------CACTCTCATGAAGGAAGAAGGAGTTCAGGTGGTGGTCGTTATCGAAACCG 3572

Qy        800 gGlnProValIleAlaProLeuAspArgTyrAlaGluArgAlaValAsnGlnValHisAl 820

Db       3573 AAGTCCGTCAGATTCTGACATGGAAGATTATTCTCCTCCTCCCAGCCTTAGTGAGGTTGC 3632

Qy        820 aArgAsnPheProSerThrIleAlaThrMetGluAlaSerLysLeuCysAspArgArgAs 840

Db       3633 TAGGAAA--------------------ATGAAGAAAAAAGAAAAACAGAAGAAAAGGAA 3671

Qy        840 nAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHisLeuLeuArgMe 860

Db       3672 AGCATATGAACCAAAACTAACACCTGAAGAA-------------------------AT 3704

Qy        860 tMetAspProSerThrLeuAlaSerPheProAsnTyrGlyThrSerSerArgAsnGlnMe 880

Db       3705 GATGGACTCTTCAACTTTTAAGAGATTC------------ACAGCCTCAATAGAGAATAT 3752

Qy        880 tGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsnGlnTyrLysGlySerThrSerTh 900

Db       3753 TTTGGATAATTTGGAAGATATGGATTTT----------------------------AC 3782

Qy        900 rSerTyrGlySerAsnLeuAsnGlyLysIleProLeuThrPheGluAspLeuSerArgHi 920

Db       3783 TGCGTTTGGTGAT------GATGATGAAATTCCTCAGGAACTGCTC---TTAGGAAAACA 3833

Qy        920 sGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGlyValLeuGlySe 940

Db       3834 TCAGCTTAATGAACTTGGCAGTGAATCTGCTAAAATAAAAGCAATGGGTATAATGGATAA 3893

Qy        940 rLeu--------------------------------LeuGlnLysGluIleAlaAsnTr 949

Db       3894 GCTTTCAACTGACAAAACTGTGAAAGTCTTAAATATCTTGGAGAAGAATATT-------- 3945

Qy        949 pSerGluAsnCysGlyThrGlnSerGlyTyrLysLeuGlyValSerThrGlyIleThrSe 969

Db       3946 ------------------CAGGATGGGTCAAAGCTTTCCACT---------TTGTTAAA 3977

Qy        969 rHisGlnMetAsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaLy 989

Db       3978 TCATAAT---AACGATACTGAAGAAGAAGAAAGGTTA---------------------- 4011

Qy        989 sTrpAsnAlaLeuGlnLeuGlySerValSerSerSerAlaAsp 1003

Db       4012 -TGGAGAGACCTTATTATGGAGAGAGTTACAAAATCAGCGGAT 4053

RESULT 4
US-11-124-368A-75
; Sequence 75, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75
;   LENGTH: 4174
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-75

Alignment Scores:
Pred. No.:              1.6e-06     Length:          4174
Score:                  202.50      Matches:         185
Percent Similarity:     34.8%       Conservative:    112
Best Local Similarity:  21.7%       Mismatches:      319
Query Match:            3.7%        Indels:          238
DB:                     12          Gaps:            38

US-09-828-068-2 (1-1057) x US-11-124-368A-75 (1-4174)

Qy         24 GlyThrGlyAlaValAlaProValLeuGluLeuThrAlaThrProArgGlnAsp------ 41

Db         62 GGGACCAGAGCAGTGCTGGGGCCGGGGCGGGTCACCGGGACTCCCCGCCAGGCCTCCCCG 121

Qy         42 ------AlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGluHisPhe 59

Db        122 CCACTCTCCGCGGCGCATTCCGGGAGGCAGCGGCCGCAGCGGCCTCGC------CATGTC 175

Qy         60 SerIleArgGlyTyrVal-AlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSe 79

Db        176 CCAGCCCGGCCAGAAGCCCGCCGCCTCCCCGCGG------CCCCGGCGAGCAGCCGCCGC 229

Qy         79 rArgIlePheHis-----------------------------------------------As 84

Db        230 CCGCCGCACCCATGAGCATGTCAGTGAAAAAACCAGTGAATCGCCTTCCAAACCAGGAGA 289

Qy         84 pGlnLysLysCysAspGluHisLysAlaSerSerSerProPheSer-------------- 99

Db        290 AAAGAAAGGATCAGATGAGAAAAAAGCAGCAAGCCTCGGCAGCAGTCAATCCTCCAGAAC 349

Qy        100 -------------ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLe 115

Db        350 CTATGCTGGTGGAACAGCCTCGGCCACCAAGGTGTCAGCTTCCTCTGGTGCAACCAGCAA 409

Qy        115 uLysThrSerAspAsn------------------------------------------- 120

Db        410 GTCTTCCAGTATGAATCCCACAGAAACCAAGGCTGTAAAAACAGAACCTGAGAAGAAGTC 469

Qy        121 ----GlyThrAlaProArgThrLeuPro------AlaLysGlnAsnGlyThrSerAspGl 137

Db        470 ACAGTCAACCAAGCCAAAAAGCCTACCCAAGCAGGCATCAGATACAGGAAGTAACGATGC 529

Qy        137 yCysSerIleThrPheVal----------------------------ArgSerThrPh 147

Db        530 TCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAGCAGCCATCAGAGAAATCAACAGA 589

Qy        147 eValProAla-----------------SerValGlySerGlnLysValSerProSerTh 161

Db        590 ACCAAAGACTAAACCACAAGACATGATTTCTGCTGGTGGAGAGAGTGTTGCTGGTATCAC 649

Qy        161 rGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLysSerValGlnGluGl 181

Db        650 TGCAATATCT---GGCAAGCCGGGTGACAAGAAAAAAGAAAAGAAATCATTAACCCCAGC 706

Qy        181 y------AsnAspSerLysCysAsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAs 199

Db        707 TGTGCCAGTTGAATCTAAACCGGATAAACCATCGGGAAAGTCAGGC---ATGGATGCTGC 763

Qy        199 nThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAs 219

Db        764 TTTGGATGACTTAATAGATACTTTAGGAGGACCTGAAGAA---------------------- 803

Qy        219 nValSerGluAspAsnThrSerValAspValGlyAlaLeuProGluVal------ProGl 237

Db        804 -ACTGAAGAAGAAAATACAACGTAT---------ACTGGACCAGAAGTTTCAGATCCAAT 853

Qy        237 nIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSe 257

Db        854 GAGTTCCACCTACATAGAGGAATTGGGTAAAAGAGAAGTCACAATTCCTCCAAAATATAG 913
```

```
Qy       257 rGluValValLeuLysArg----------------------------------------- 263
             |||:::::::   |||:::
Db       914 GGAACTATTGGCTAAAAAGGAAGGGATCACAGGGCCTCCTGCAGACTCTTCGAAACCCAT 973

Qy       263 ------------------------------------------------------------ 263

Db       974 AGGGCCAGATGATGCTATAGACGCCTTGTCATCTGACTTCACCTGTGGGTCGCCTACAGC 1033

Qy       264 -------AsnGluAspGluAsnGlyLysThrGluGluThrLeuValAlaGluGlnCysAs 281
                     :::   |||      ::::::   |||   |||   |||:::

Db      1034 TGCTGGAAAGAAAACTGAAAAAGAGGAATCTACAGAAGTTTTAAAAGCTCAGTCAGCAGG 1093

Qy       281 nLeuThrLysAspProAsnProMetSerGlyLysGluArg-------------------- 294
              :::          |||       |||:::|||

Db      1094 GACAGTCAGAAGTGCTGCTCCACCCCAAGAGAAGAAAAGAAAGGTGGAGAAGGATACAAT 1153

Qy       295 ----AspGlnValAlaGluGln-CysAsnLeuThrLysAspProLysProValSerGlyG 313
                 ||||||      |||    |||   |||          |||             |||:

Db      1154 GAGTGATCAAGCACTCGAGGCTCTGTCGGCTTCACTGGGCACCC------------GGCA 1201

Qy       313 lnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerL 333
              :::::          |||                  |||::::::|||    :

Db      1202 AGCAG--------------AACCTGAGCTCGACCTCCGCTCAATTAAGGAAGTCGATGA 1246

Qy       333 ysSerLysArgLysThrAspLysLys-----------------LeuMetLysLysGlnG 347
              ::   |||   |||          ||||||                |||   ||||||

Db      1247 GGCAAAAGCTAAAG------AAGAAAAACTAGAGAAGTGTGGTGAGGATGATGAAACAAT 1300

Qy       347 lnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgL 367
               |||    :::   |||          |||          |||   |||          :

Db      1301 CCCATCT-GAGTACAGATTAAAACCAGCCACGGATAAAGATGGAAAACCACTATTGCCAG 1359

Qy       367 ysProLys---------LysValArgLeuLeuSerGluIleIleAsnAlaAsnGlnValG 384
              ::|||:::          |||   |||       ||||||:::|||:::          :

Db      1360 AGCCTGAAGAAAAACCCAAGCCTCGGAGTGAATCAGAACTCATTGATGAACTTTCAGAAG 1419

Qy       384 luAspSerArgSerAspGluValHisArgGluAsnAlaAlaAspPro---------CysG 401
              ::      ||||||:::         :::|||      :::   |||          |

Db      1420 ATTTTGACCGGTCTGAA------TGTAAAGAGAAACCATCTAAGCCAACTGAAAAGACAG 1473

Qy       401 luAspAspArgSerThrIleProValProMetGluValSerMet---------------- 415
              ||:::   ::::::       |||   |||:::      ::::::

Db      1474 AAGAATCTAAGGCCGCTGCTCCAGCTCCTGTGTCGGAGGCTGTGTGTCGGACCTCCATGT 1533

Qy       416 ----------------------AspIleProValSerAsnHisThrValGlyGluAspG 428
                                    :::                 ||||||   :::|||

Db      1534 GTAGTATACAGTCAGCACCCCCTGAGCCGGCTACCTTGAAGGGCACAGTGCCAGATGATG 1593

Qy       428 lyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGlyS 448
               :::::::::      :::         :::|||   :::|||      :::||||||

Db      1594 CTGTAGAAGCCTTGGCTGATAGCCTGGGGAAAAAGGAAGCAGAT---CCAGAAGATGGAA 1650

Qy       448 erSerLeuMetAsnTrpLeuAsnGlyLysLysLys--------------ArgThrGlyS 463
                  :::|||:::   :::       |||   |||                :::   |||

Db      1651 AACCTGTGATGGATAAAGTCAAGGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCTTGGTG 1710

Qy       463 erValHisHisThrVal--------------AlaHisProAlaGlyAsnLeuSerAsnL 478
                       |||:::                        :::          |

Db      1711 AAAAAGAAGAAACAATTCCTCCTGATTATAGATTAGAAGAGGTCAAGGATAAAGATGGAA 1770

Qy       478 ysLysValThrProThrAlaSerThrGlnHis---------------------------- 488
              ||   :::   |||       |||   :::

Db      1771 AGCCACTCCTGCCAAAAGAGTCTAAGGAACAGCTTCCACCCATGAGTGAAGACTTCCTTC 1830

Qy       489 --------------AspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHisL 504
                           :::|||   :::       :::|||

Db      1831 TGGATGCTTTGTCTGAGGACTTCTCTGGTCCACAAAATGCTTCATCTCTTAAATTTGAAG 1890

Qy       504 ysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGlyL 524
                    :::         :::||||||:::   :::|||

Db      1891 ATGCTAAACTTGCTGCTGCCATCTCTGAAGTGGTTTCCCAA------------------- 1931

Qy       524 ysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGlyG 544
```

```
                                          :::   |||||||||          |||
Db      1932 ------------------------------ACCCCAGCTTCAACGACCCAAGCTGGAG 1959

Qy       544 luSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMetG 564
                    |||:::                  :::::::::          ::::

Db      1960 CCCCACCCCGTGAT--------------------ACCTCGCAGAGTGACAAAGACCTCG 1998

Qy       564 luThrGluAsnSerValLeuSerHisSer---AlaLysValSerProAlaGluHisAspI 583
              ::                |||||   |||      :::      |||          :::

Db      1999 ATGATGCCTTGGATAAACTCTCTGACAGTCTAGGACAAAGGCAGCCTGACCCAGATGAGA 2058

Qy       583 leGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysLysGlnLysL 603
               :::   |||   |||                      |||   |||:::|||   :::

Db      2059 ACAAACCAATGGAAGAT--------------------AAAGTAAAGGAAAAAGCTAAAG 2097

Qy       603 euGluValThrArgGluLysGlnThrMetIleAspAsp---IleProMetAspIleValG 622
               |||     |||:::|||              ||||||   ||||||   :::

Db      2098 CTGAA---CATAGAGACAAGCTTGGAGAAAGAGATGACACTATCCCACCTGAATACAGAC 2154

Qy       622 luLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspI 642
               ||||||      |||      ::::::   :::          |||      |||:::

Db      2155 ATCTCCTGGATGATAATGGACAGGACAAACCAGTGAAGCCACCTACAAAGAAATCAGAG- 2213

Qy       642 leAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysA 662
                          ||||||      ||||||||          |||   |||

Db      2214 ----------GATTCAAAGAAACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCAG 2262

Qy       662 spGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------------ 677
                       |||          |||      :::|||:::

Db      2263 GAGATCTGGACAGCTGTCCCTCCACTACAGAAACCTCACAGAACACAGCAAAGGATAAGT 2322

Qy       678 --LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeu---------A 694
               |||      ||||||   |||         :::      |||                |

Db      2323 GCAAGAAGGCTGCTTCCAGCTCCAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAG 2382

Qy       694 laLeuThrThrGlnGluSerProHisProGlnAsn 705
              ||   ||||||:::|||:::      |||::::::
Db      2383 CAAAGACAACAGAGGAAACTTCCAAGCCAAAAGAT 2417
```

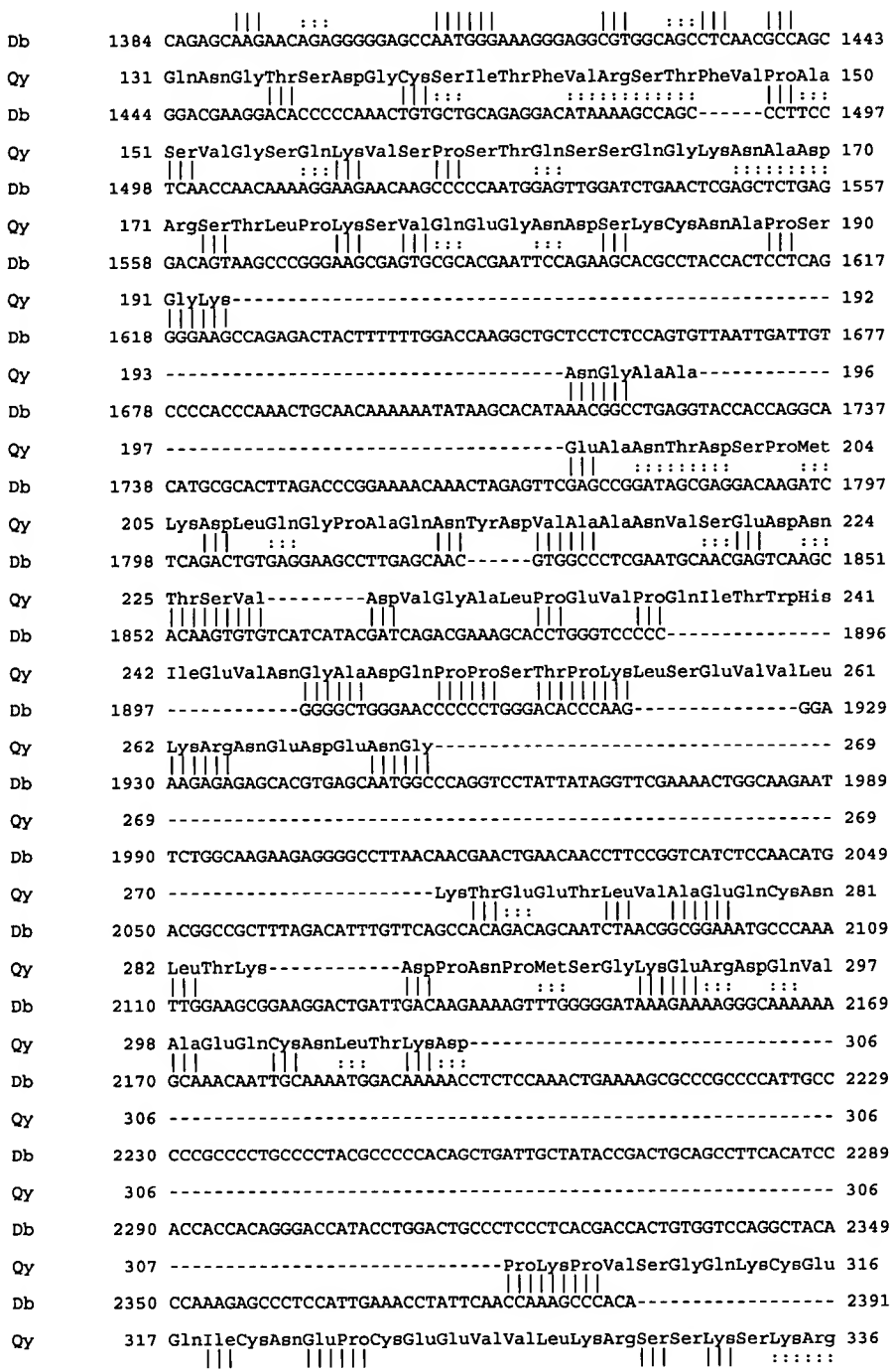
RESULT 5
US-10-330-773-84
; Sequence 84, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 5545
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-330-773-84

Alignment Scores:

Pred. No.:	2.43e-06	Length:	5545
Score:	202.00	Matches:	221
Percent Similarity:	30.6%	Conservative:	147
Best Local Similarity:	18.4%	Mismatches:	398
Query Match:	3.7%	Indels:	436
DB:	7	Gaps:	52

US-09-828-068-2 (1-1057) x US-10-330-773-84 (1-5545)

```
Qy        93 AlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeu 112
             |||::::::::::|||   |||         |||              |||:::   |||

Db      1333 GCCGCGGCTGCCCCAGGCTCCGAGGTCAGCTTCACCGAG---------TCCAGAGGGCTG 1383

Qy       113 ---AspLysLeuLysThrSerAspAsnGlyThrAlaProArg---ThrLeuProAlaLys 130
```


Db 1384 CAGAGCAAGAACAGAGGGGGAGCCAATGGGAAAGGGAGGCGTGGCAGCCTCAACGCCAGC 1443
Qy 131 GlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheValProAla 150
Db 1444 GGACGAAGGACACCCCCAAACTGTGCTGCAGAGGACATAAAAGCCAGC------CCTTCC 1497
Qy 151 SerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAsp 170
Db 1498 TCAACCAACAAAAGGAAGAACAAGCCCCCAATGGAGTTGGATCTGAACTCGAGCTCTGAG 1557
Qy 171 ArgSerThrLeuProLysSerValGlnGluGlyAsnAspSerLysCysAsnAlaProSer 190
Db 1558 GACAGTAAGCCCGGGAAGCGAGTGCGCACGAATTCCAGAAGCACGCCTACCACTCCTCAG 1617
Qy 191 GlyLys---------- 192
Db 1618 GGGAAGCCAGAGACTACTTTTTTGGACCAAGGCTGCTCCTCTCCAGTGTTAATTGATTGT 1677
Qy 193 ----------AsnGlyAlaAla---------- 196
Db 1678 CCCCACCCAAACTGCAACAAAAAATATAAGCACATAAACGGCCTGAGGTACCACCAGGCA 1737
Qy 197 ----------GluAlaAsnThrAspSerProMet 204
Db 1738 CATGCGCACTTAGACCCGGAAAACAAACTAGAGTTCGAGCCGGATAGCGAGGACAAGATC 1797
Qy 205 LysAspLeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsn 224
Db 1798 TCAGACTGTGAGGAAGCCTTGAGCAAC------GTGGCCCTCGAATGCAACGAGTCAAGC 1851
Qy 225 ThrSerVal---------AspValGlyAlaLeuProGluValProGlnIleThrTrpHis 241
Db 1852 ACAAGTGTGTCATCATACGATCAGACGAAAGCACCTGGGTCCCCC---------- 1896
Qy 242 IleGluValAsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeu 261
Db 1897 ----------GGGGCTGGGAACCCCCCTGGGACACCCAAG----------GGA 1929
Qy 262 LysArgAsnGluAspGluAsnGly---------- 269
Db 1930 AAGAGAGAGCACGTGAGCAATGGCCCAGGTCCTATTATAGGTTCGAAAACTGGCAAGAAT 1989
Qy 269 ---------- 269
Db 1990 TCTGGCAAGAAGAGGGGCCTTAACAACGAACTGAACAACCTTCCGGTCATCTCCAACATG 2049
Qy 270 ----------LysThrGluGluThrLeuValAlaGluGlnCysAsn 281
Db 2050 ACGGCCGCTTTAGACATTTGTTCAGCCACAGACAGCAATCTAACGGCGGAAATGCCCAAA 2109
Qy 282 LeuThrLys----------AspProAsnProMetSerGlyLysGluArgAspGlnVal 297
Db 2110 TTGGAAGCGGAAGGACTGATTGACAAGAAAAGTTTGGGGGATAAAGAAAAGGGCAAAAAA 2169
Qy 298 AlaGluGlnCysAsnLeuThrLysAsp---------- 306
Db 2170 GCAAACAATTGCAAAATGGACAAAAACCTCTCCAAACTGAAAAGCGCCCGCCCCATTGCC 2229
Qy 306 ---------- 306
Db 2230 CCCGCCCCTGCCCCTACGCCCCCACAGCTGATTGCTATACCGACTGCAGCCTTCACATCC 2289
Qy 306 ---------- 306
Db 2290 ACCACCACAGGGACCATACCTGGACTGCCCTCCCTCACGACCACTGTGGTCCAGGCTACA 2349
Qy 307 ----------ProLysProValSerGlyGlnLysCysGlu 316
Db 2350 CCAAAGAGCCCTCCATTGAAACCTATTCAACCAAAGCCCACA---------- 2391
Qy 317 GlnIleCysAsnGluProCysGluGluValValLeuLysArgSerSerLysSerLysArg 336
Db 2392 ---ATTATGGGGGAGCCCATCACCGTGAACCCAGCTCTGGTGTCACTCAAAGACAGAAAG 2448
Qy 337 LysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAla 356
Db 2449 AAAAAGGAGAAGCGAAAGCTCAAGGACAAAGAAGGGAAAGAGACTGGAGGTCCCAAAATG 2508
Qy 357 AspValSerAspAlaLysLeuCysArgArgLysProLysLysValArgLeuLeuSerGlu 376
Db 2509 GATGCTAAGCTGGGGAAGCTGGAGGAAGCCAAGGCGGCTAGCAAGGATTTGTCTGGGCAT 2568
Qy 377 IleIleAsn----------AlaAsnGlnValGluAspSerArg 387
Db 2569 TTTTTAAAGGACCATCTCGGTAAGAGCGAAGGGCTGGCTAATGGATTGTCAGAGTCTCAG 2628
Qy 388 ----------SerAspGluValHis------ArgGluAsn 395
Db 2629 GAGAGCCGCATGGCCAGCATCAAAGCCGAGGCCGATAAGGTTTACACATTCACGGACAAT 2688
Qy 396 AlaAlaAspProCysGluAspAspArgSerThrIleProValProMetGluValSerMet 415
Db 2689 GCCCCCAGCCCTTCCATCGGGAGTGCCTCACGCATGGAATGCAGTACCTTGGTGAATGGG 2748
Qy 416 AspIleProValSerAsn---HisThrValGlyGluAspGlyLeuLysSerSerLysAsn 434
Db 2749 CAAGCACCCATGGCACCACTGCATGTATTGACCCAGAATGGGGCTGAGAGCGCAGCAGCT 2808
Qy 435 LysThrLysArg---LysTyrSerAspVal---------ValAspAspGlySerSerLeu 450
Db 2809 AAGACCAGCAGTCCAGCCTATTCGGACATATCAGATGCTGCAGATGATGGC--------- 2859
Qy 451 MetAsnTrpLeuAsnGlyLysLysLysArgThrGlySerValHisHisThrValAlaHis 470
Db 2860 ----------GGTTCTGACAGCAGGTCAGAGGGCATGAGGTCAAAGGCCAGTTCC 2904
Qy 471 ProAlaGlyAsnLeuSerAsnLys----------LysValThrProThrAlaSerThr 486
Db 2905 CCATCGGATACCTTTTCTAACAAAGACGGTGTTGTGAAAGGGCATCCTTCAACTTCAGCA 2964
Qy 487 GlnHisAspAspGluAsnAspThrGlu----------AsnGlyLeuAspThrAsnMet 502
Db 2965 CAGCCATCTCAGCTGAAAGAGTCCCATTCTCCCTATTACCATGGCTATGAGCCT------ 3018
Qy 503 HisLysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLys 522
Db 3019 TACTATTCTCCGAGTTACATGCACCCTGGACAGGTTGGTGCCCCGGCAGCTGGGAATGGT 3078
Qy 523 GlyLysThrAlaGlyLeuSer---LysGlyLysThrHisSerAlaAlaSerThrLysTyr 541
Db 3079 GGGAGCACACAGGGCATGAAGATCAAGAAGGAGTCAGAGGAAGATGCGGAGAAGAAAGAC 3138
Qy 542 GlyGlyGluSerThrArgAsnGlyGlnAsnIleHisVal---------LeuSerAlaGlu 558
Db 3139 AAGGCGGAACAGTTGGAGTCCAAGAAAGTAGACCATACTTCTGCACCCCTGCAGCCTCAG 3198
Qy 559 AspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSer---------- 573
Db 3199 CACCAGTCAGTCATCACCCAGAGACACCCTGCCCTGGCTCAGTCACTTTATTACGGCCAG 3258
Qy 574 ----------AlaLysValSerProAlaGluHis 581
Db 3259 TATGCCTACGGGCTCTATATGGACCAGAAATCTCTGATGGCTACCAGCCCTGCCTACAGA 3318
Qy 582 AspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnGln 601
Db 3319 CAG---CAGTATGAGAAGTACTATGAGGACCAGAGACTGGCAGAGCAGAAAATGGCCCAG 3375
Qy 602 LysLeuGluValThrArgGluLysGlnThrMetIleAspAspIleProMetAspIleVal 621
Db 3376 AGTGGGAGAGGAGACTGTGAAAGAAAGGCT---------GAACTCCCCTTG--------- 3417
Qy 622 GluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp 641
Db 3418 AAAGAACTGGGCAAGGAAGACAATAAGCAGAAAAACATGCCATCAGCCACTATCTCA--- 3474

```
Qy       642 IleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLys 661
Db      3475 ------------AAAGCTCCCTCTACCCCGGAGCCTAAC-----------------AAA 3504

Qy       662 AspGlySerAspTyrAlaSerSerValPheAsp---ThrAsnSerGlnGlnLysSerLeu 680
Db      3505 AACCATTCTAAACTAGGGCCATCAGTGCCTAATAAAACTGAGGAGACAGGTAAATCACAG 3564

Qy       681 AlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThrThrGlnGluSer 700
Db      3565 CTTCTCTCCAGTCACCAGCAGCAGCTTCAGGCC-------------------------- 3597

Qy       701 ProHisProGlnAsnPheGlnSerThrGlnGluGlnGlnThrHisLeuArgMetGlu--- 719
Db      3598 ---------GACAGCTTCAAAGCTAAGCAGATGGAAAACCACCAGCTTATTAAGGAGGCT 3648

Qy       720 ---GluMetValThrIleAlaAlaSerSerProLeuPheSerHisHisAspAspGlnTyr 738
Db      3649 GTAGAAATGAAGTCTGTCATGGACTCTATG----------------------AAGCAG 3684

Qy       739 IleAlaGluAlaProThrGluHisTrpGlyArgLysAspAlaLysLysLeuThrTrpGlu 758
Db      3685 ACAGGTGTAGACCCAACCTCGAGATTT---AAACAAGATCCAGAGTCAAGGACATGGCAC 3741

Qy       759 GlnPhe---------------------------------------------------- 760
Db      3742 CATTATGTATACCAGCCCAAGTACCTGGACCAACAAAAGCCAGAAGAACTTGACAGAGAA 3801

Qy       761 LysAlaThrThrArgAsnSerProAlaAlaThrCysGlyAlaGlnPhe----------- 776
Db      3802 AAGAAATTGAAAGAGGACAGTCCAAGGAAGACTCCTAACAAAGAGAGTGGCGTGTCCAGT 3861

Qy       777 ----------------------------------------------ArgProGlyIle 780
Db      3862 CTCCCCGTGTCATTAACAAACATCAAAGAGGAGCCCAAAGAGGGCAAGCGTCCTGATTCC 3921

Qy       781 GlnAlaVal--------------------AspLeuThrSerThrHisValMetGlySer 793
Db      3922 CAGTCGGTGGAGGAGAACAAGCTAAAGAATGATGATCGGAAGACACCGGTGAACTGGAAG 3981

Qy       794 SerSerAsnTyrAlaSerArgGlnProValIleAlaProLeuAspArgTyrAlaGluArg 813
Db      3982 GACTCTCGG---GGGACACGAGTGGCCGTGTCCTCACCCATGAGTCAGCAC---CAGTCC 4035

Qy       814 AlaValAsnGlnValHisAlaArgAsnPheProSerThrIleAlaThrMetGluAlaSer 833
Db      4036 TACATACAATACTTGCATGCT------TATCCT-------------------------- 4062

Qy       834 LysLeuCysAspArgArgAsnAlaGlyGlnValValLeuTyrProLysGluSerMetPro 853
Db      4063 ------------------------------------TACCCACAGATGTATGACCCC 4083

Qy       854 AlaThrHisLeuLeuArgMetMetAspProSerThrLeuAlaSerPheProAsn------ 871
Db      4084 AGCCACCCTGCATACCGGGCTGTCTCTCCTGTCTTAATGCACAGTTACCCTGGGGCCTAT 4143

Qy       872 --------------------------TyrGlyThrSerSerArgAsnGlnMetGluSer 882
Db      4144 CTCTCTCCAGGATTTCATTATCCTGTTTATGGAAAGATGTCTGGGAGAGAAGAGGCAGAG 4203

Qy       883 GlnLeuHisAsnSer---------------------------------------------- 887
Db      4204 AAAGTCAATACTAGTCCCAGCATCAACACAAAAACAGCCAGCGAAGCCAAAGCCCTGGAC 4263

Qy       888 ------GlnTyrAlaHisAsnGlnTyr--------------------------LysGly 896
Db      4264 CTACTCCAACATCACGCCAACCAATACCGCAGCAAGTCCCCTGCTCCTGTGGAGAAGGCT 4323

Qy       897 SerThrSerThrSerTyrGlySerAsnLeuAsnGlyLysIleProLeu-ThrPheGluAs 916
Db      4324 TCCACGGAGCGAGAACG-GGAAGCAGAGCGGGAACGGGATCGCCACTCCCCCTTC----- 4377

Qy       916 pLeuSerArgHisGlnLeuHisAspLeuHisArgProLeuArgProHisProArgValGl 936
Db      4378 ----AGCCAGCGGCACCTGCACACACACCAC--------------CACACCCATGTTGG 4418

Qy       936 yValLeuGlySerLeuLeuGlnLysGluIleAlaAsnTrpSerGluAsnCysGlyThrGl 956
Db      4419 CATG---------------------------------------------------------- 4422

Qy       956 nSerGlyTyrLysLeu---------------GlyValSerThrGlyIleThrSerHisGl 971
Db      4423 ----GGCTACCCTCTCATCCCTGGTCAATACGACCCTTTTCAAGGCTTGACCTCTGCTGC 4478

Qy       971 nMet------AsnArgLysGluHisPheGluAlaLeuAsnSerGlyMetPheSerAlaLy 989
Db      4479 CCTAGTTGCCTCTCAGCAGGTGGCTGCCCAGGCATCTGCTTCTGGAATGTTTCCTGCACA 4538

Qy       989 s 989
Db      4539 G 4539
```

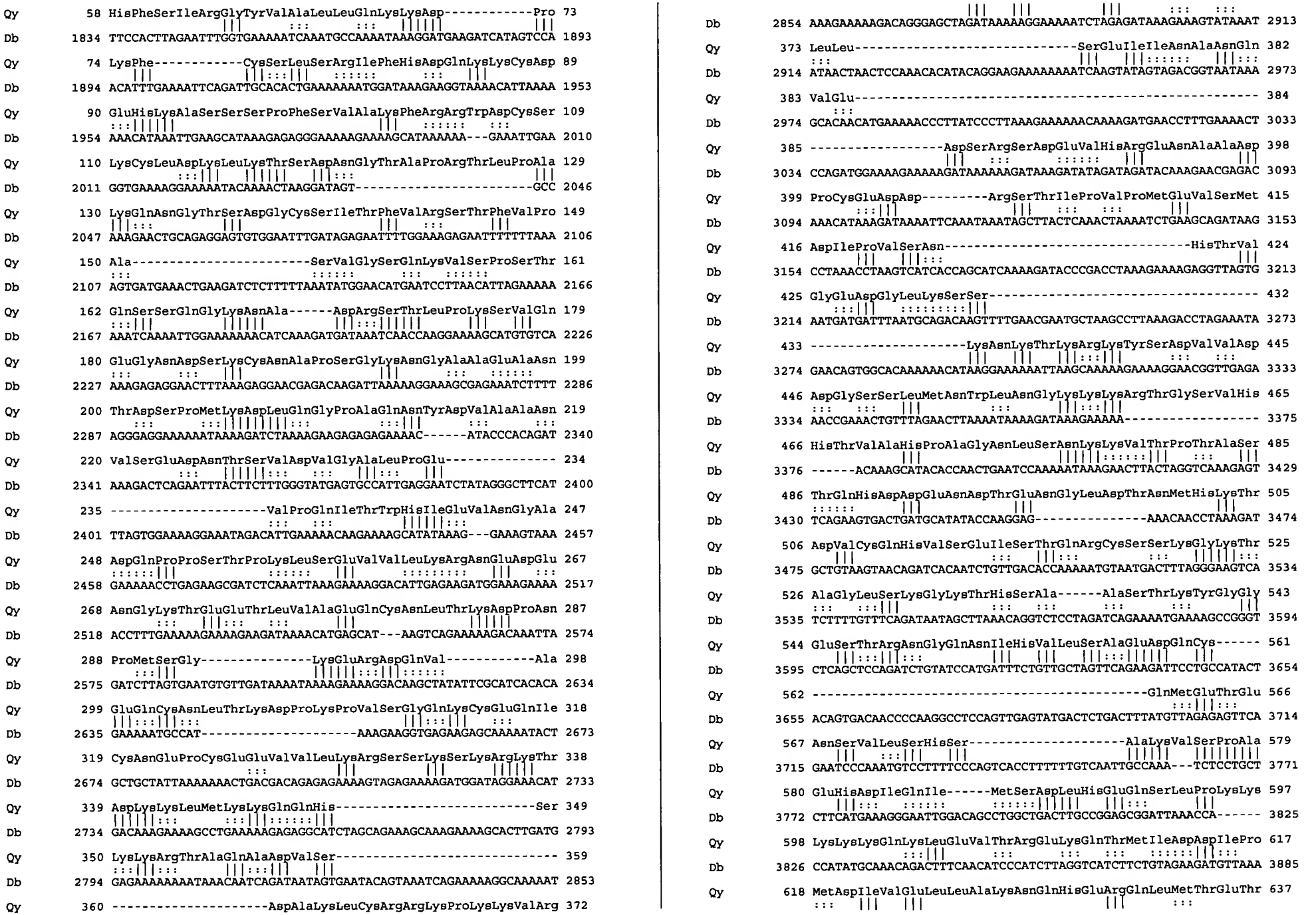
RESULT 6
US-10-909-125-825
; Sequence 825, Application US/10909125
; Publication No. US20050261218A1
; GENERAL INFORMATION:
; APPLICANT: Esau, Christine
; APPLICANT: Lollo, Bridget
; APPLICANT: Bennett, C. Frank
; APPLICANT: Freier, Susan M.
; APPLICANT: Griffey, Richard H.
; APPLICANT: Baker, Brenda F.
; APPLICANT: Vickers, Timothy
; APPLICANT: Marcusson, Eric G.
; APPLICANT: Koller, Erich
; APPLICANT: Swayze, Eric
; APPLICANT: Jain, Ravi
; APPLICANT: Bhat, Balkrishen
; APPLICANT: Peralta, Eigen
; TITLE OF INVENTION: Oligomeric Compounds And Compositions For Use In Modulation
; TITLE OF INVENTION: Of Small Non-Coding RNAs
; FILE REFERENCE: ISIS0080-100 (CORE0016US)
; CURRENT APPLICATION NUMBER: US/10/909,125
; CURRENT FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: US 60/492,056
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/516,303
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: US 60/531,596
; PRIOR FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/562,417
; PRIOR FILING DATE: 2004-04-14
; NUMBER OF SEQ ID NOS: 2184
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 825
; LENGTH: 6189
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-909-125-825

Alignment Scores:

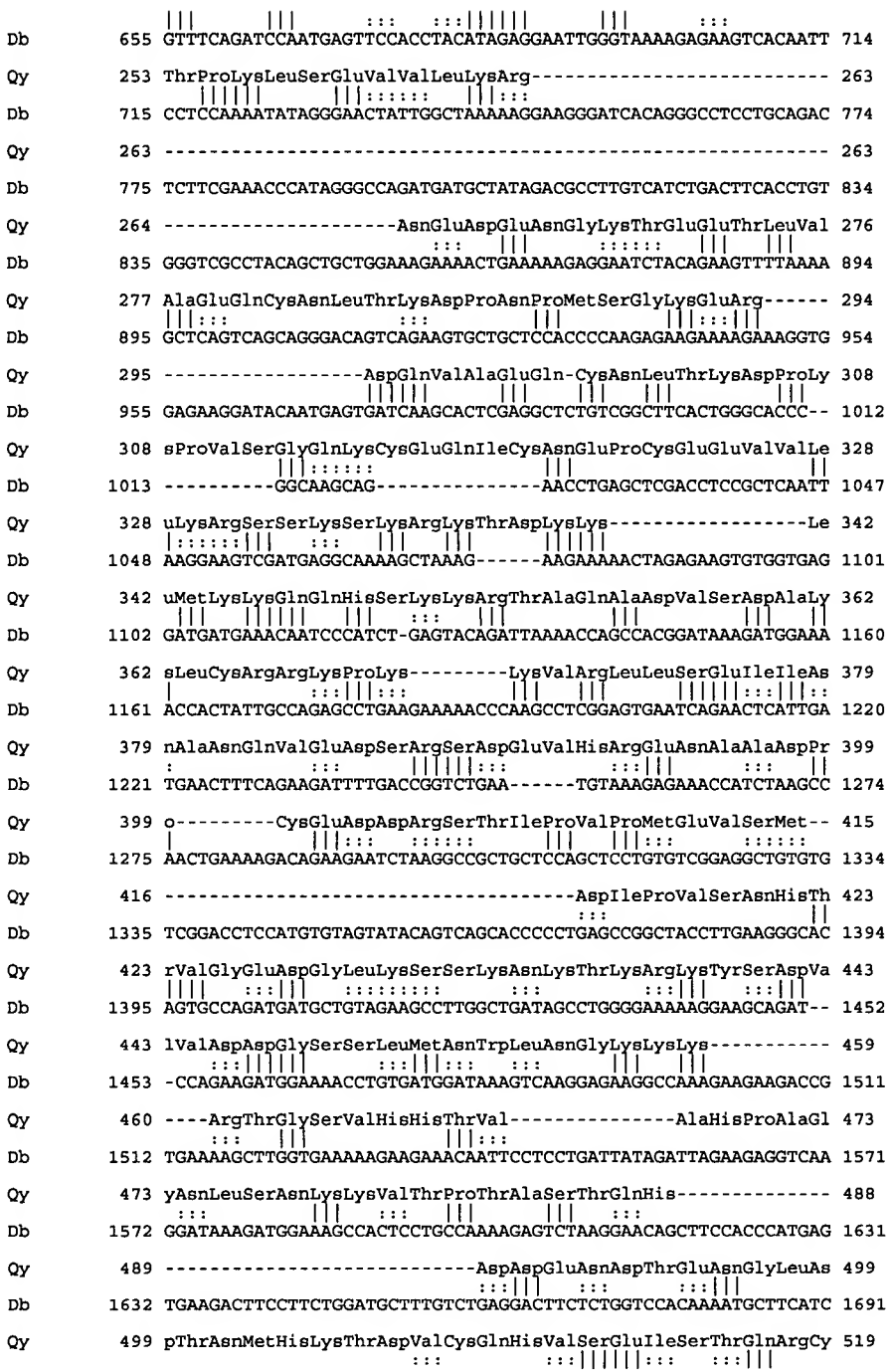
Pred. No.:	2.74e-06	Length:	6189
Score:	202.00	Matches:	192
Percent Similarity:	33.8%	Conservative:	161
Best Local Similarity:	18.4%	Mismatches:	391
Query Match:	3.7%	Indels:	300
DB:	8	Gaps:	43

US-09-828-068-2 (1-1057) x US-10-909-125-825 (1-6189)

```
Qy        38 ProArgGlnAspAlaAlaAlaGluAlaGlyValAspGluProAlaGlnHisGlnCysGlu 57
Db      1774 CCAGAAAGTTCAAGTGTAAAATCTTGTAAGCATAAGGAAAAAAGCAAACATCAGAAAGAT 1833
```



```
Db       3886 CTAATTATAAGCGAG--------------------GGGAGACCTACCATAGAAGTTCGA 3924

Qy        638 AspCysSerAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIle 657

Db       3925 AGATGTAGC---------------------------------------------------- 3933

Qy        658 ValAlaAlaLysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln 677

Db       3934 ------------------------ATGCCTTCTGTCATTTGTGAACATACCAAACAATTC 3969

Qy        678 LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeuAlaLeuThr--- 696

Db       3970 CAAACAATATCAGAAGAGAGCAATCAAGGTAGCTTA--------------TTAACTGTG 4014

Qy        697 ---------ThrGlnGluSerProHisProGlnAsnPheGlnSerThrGlnGluGln--- 712

Db       4015 CCAGGAGATACTAGTCCTTCTCCCAAACCTGAGGTATTCTCAAATGTGCCTGAAAGAGAC 4074

Qy        713 --------------GlnThrHisLeuArgMetGluGluMetValThrIleAlaAlaSer 727

Db       4075 CTTTCAAATGTATCTAACATACATTCCAGTTTTGCAACTTCTCCAACTGGAGCTTCAAAC 4134

Qy        728 SerProLeuPheSerHisHisAspAspGlnTyrIleAlaGluAlaProThrGluHisTrp 747

Db       4135 AGCAAGTATGTTTCAGCTGATAGAAATCTCATCAAGAATACTGCCCCAGTGAAC------ 4188

Qy        748 GlyArgLysAspAlaLysLysLeuThrTrpGluGlnPheLysAlaThrThrArgAsnSer 767

Db       4189 ----------------------------------------ACTGTAATGGACAGT 4203

Qy        768 ProAlaAlaThrCysGlyAlaGlnPheArgProGlyIleGlnAlaValAspLeuThrSer 787

Db       4204 CCA-------------------------------------GTGCATTTAGAGCCA 4221

Qy        788 ThrHisValMetGlySerSerSerAsnTyrAlaSerArgGlnProValIleAlaProLeu 807

Db       4222 TCTAGTCAGGTTGGTGTGATCCAGAATAAATCATGGGAGATGCCTGTT------------ 4269

Qy        808 AspArgTyrAlaGluArgAlaValAsnGlnValHisAlaArgAsnPhe---------Pro 824

Db       4270 GATAGA--------------CTAGAGACATTAAGCACCAGAGACTTTATCTGCCCAAAT 4314

Qy        825 SerThrIleAlaThrMetGluAlaSer------LysLeuCysAsp--------------- 837

Db       4315 TCTAACATACCTGATCAAGAATCCTCTCTTCAGAGTTTTTGTAATTCTGAAAATAAGGTA 4374

Qy        838 ---ArgArgAsnAlaGlyGlnValValLeuTyrProLysGluSerMetProAlaThrHis 856

Db       4375 TTGAAAGAAAATGCTGATTTTTTATCCCTGCGCCAGACTGAA---CTGCCAGGAAACTCT 4431

Qy        857 LeuLeuArg------------MetMetAspProSerThrLeuAlaSerPheProAsnTyr 872

Db       4432 TGTGCTCAGGATCCGGCATCCTTTATGCCTCCACAGCAGCCTTGCTCTTTCCCCAGCCAA 4491

Qy        873 GlyThrSerSerArgAsnGlnMetGluSerGlnLeuHisAsnSerGlnTyrAlaHisAsn 892

Db       4492 TCACTTTCAGATGCTGAATCGATTTCTAAACATATGTCTTTGTCA---TATGTTGCTAAT 4548

Qy        893 GlnTyrLysGly 896

Db       4549 CAAGAGCCAGGT 4560

RESULT 7
US-11-124-368A-70
; Sequence 70, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 3989
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-70

Alignment Scores:
Pred. No.:              2.23e-06        Length:         3989
Score:                  200.50          Matches:        178
Percent Similarity:     35.1%           Conservative:   109
Best Local Similarity:  21.8%           Mismatches:     310
Query Match:            3.6%            Indels:         221
DB:                     12              Gaps:           38

US-09-828-068-2 (1-1057) x US-11-124-368A-70 (1-3989)

Qy          35 ThrAlaThrProArgGlnAsp--------------------AlaAlaAlaGluAlaGly 47

Db           7 ACCGCCCCGCCCCGTCGCACTCAGAGAGCTGGGCTGGAGCTCCAAGCGGAAACCGCGGGA 66

Qy          48 ValAsp---------------GluProAlaGlnHisGlnCysGluHisPheSerIleArg 62

Db          67 GCCGAGCCCAGCTAGGAATGCAGACCTCCTGAAAACCAAGCCGAGGACTGCGGGGTCCGG 126

Qy          63 ------GlyTyrValAlaLeuLeuGlnLysLysAspProLysPheCysSerLeuSerArg 80

Db         127 TGTCCACGCAGAGTGTCAGCTTCCTCTGGTGCAACCAGCAAGTCTTCCAGTATGAATCCC 186

Qy          81 IlePheHisAspGlnLysLysCysAsp---GluHisLysAlaSerSerSerProPheSer 99

Db         187 ACAGAAACCAAGGCTGTAAAAACAGAACCTGAGAAGAAGTCACAGTCAACCAAGCTGTCT 246

Qy         100 ValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeuLysThrSerAsp 119

Db         247 GTGGTT------------------------------------CATGAGAAAAAATCCCAA 270

Qy         120 AsnGly---------------ThrAlaProArgThrLeuPro------AlaLysGlnAsn 132

Db         271 GAAGGAAAGCCAAAAGAACACACAGAGCCAAAAAGCCTACCCAAGCAGGCATCAGATACA 330

Qy         133 GlyThrSerAspGlyCysSerIleThrPheVal--------------------------- 143

Db         331 GGAAGTAACGATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAGCAGCCATCA 390

Qy         144 ---ArgSerThrPheValProAla-----------------SerValGlySerGlnLys 156

Db         391 GAGAAATCAACAGAACCAAAGACTAAACCACAAGACATGATTTCTGCTGGTGGAGAGAGT 450

Qy         157 ValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuProLys 176

Db         451 GTTGCTGGTATCACTGCAATATCT---GGCAAGCCGGGTGACAAGAAAAAAGAAAAGAAA 507

Qy         177 SerValGlnGluGly------AsnAspSerLysCysAsnAlaProSerGlyLysAsnGly 194

Db         508 TCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCGGATAAACCATCGGGAAAGTCAGGC 567

Qy         195 AlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsnTyr 214

Db         568 ---ATGGATGCTGCTTTGGATGACTTAATAGATACTTTAGGAGGACCTGAAGAA------ 618

Qy         215 AspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuProGlu 234

Db         619 --------------ACTGAAGAAGAAAATACAACGTAT---------ACTGGACCAGAA 654

Qy         235 Val------ProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProProSer 252
```



RESULT 8
US-11-124-368A-79
; Sequence 79, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 4143
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-79

Alignment Scores:

Pred. No.:	2.33e-06	Length:	4143
Score:	200.50	Matches:	186
Percent Similarity:	35.3%	Conservative:	114
Best Local Similarity:	21.9%	Mismatches:	318
Query Match:	3.6%	Indels:	232


DB: 12 Gaps: 40
US-09-828-068-2 (1-1057) x US-11-124-368A-79 (1-4143)

```
Db       1980 ATAAACTCTCTGACAGTCTAGGACAAAGGCAGCCTGACCCAGATGAGAACAAACCAATGG 2039
Qy        587 erAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrA 607
Db       2040 AAGAT--------------------AAAGTAAAGGAAAAAGCTAAAGCTGAA---CATA 2075
Qy        607 rgGluLysGlnThrMetIleAspAsp---IleProMetAspIleValGluLeuLeuAlaL 626
Db       2076 GAGACAAGCTTGGAGAAAGAGATGACACTATCCCACCTGAATACAGACATCTCCTGGATG 2135
Qy        626 ysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleG 646
Db       2136 ATAATGGACAGGACAAACCAGTGAAGCCACCTACAAAGAAATCAGAG------------G 2183
Qy        646 lnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspT 666
Db       2184 ATTCAAAGAAACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCAGGAGATCTGGACA 2243
Qy        666 yrAlaSerSerValPheAspThrAsnSerGlnGln--------------LysSerLeuA 681
Db       2244 GCTGTCCCTCCACTACAGAAACCTCACAGAACACAGCAAAGGATAAGTGCAAGAAGGCTG 2303
Qy        681 laSerGlnSerThrGlnLysGluLeuGlnGlyHisLeu---------AlaLeuThrThrG 698
Db       2304 CTTCCAGCTCCAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAGCAAAGACAACAG 2363
Qy        698 lnGluSerProHisProGlnAsn 705
Db       2364 AGGAAACTTCCAAGCCAAAAGAT 2386

RESULT 9
US-11-124-368A-81
; Sequence 81, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
;   LENGTH: 2782
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-81

Alignment Scores:
Pred. No.:              1.99e-06        Length:         2782
Score:                  199.00          Matches:        170
Percent Similarity:     34.4%           Conservative:   107
Best Local Similarity:  21.1%           Mismatches:     313
Query Match:            3.6%            Indels:         216
DB:                     12              Gaps:           36

US-09-828-068-2 (1-1057) x US-11-124-368A-81 (1-2782)

Qy         35 ThrAlaThrProArgGlnAsp--------------------AlaAlaAlaGluAlaGly 47
Db          7 ACCGCCCCGCCCCGTCGCACTCAGAGAGCTGGGCTGGAGCTCCAAGCGGAAACCGCGGGA 66
Qy         48 ValAsp--------------GluProAlaGlnHisGlnCysGluHisPheSerIleArg 62
Db         67 GCCGAGCCCAGCTAGGAATGCAGACCTCCTGAAAACCAAGCCGAGGACTGCGGGGTCCGG 126

Qy         63 --------------------------GlyTyrValAlaLeuLeuGlnLysLysAspPro 73
Db        127 TGTCCACGCAGAGTGTCAGCTTCCTCTGGTGCAACCAGCAAGTCTTCCAGTATGAATCCC 186
Qy         74 LysPheCysSerLeuSerArgIlePheHisAspGlnLysLysCysAspGluHisLysAla 93
Db        187 ACAGAAACCAAGGCTGTAAAAACA---GAACCTGAGAAGAAGTCACAGTCAACCAAGCCA 243
Qy         94 SerSerSerProPheSerValAlaLysPheArgArgTrpAspCys------SerLysCys 111
Db        244 AAAAGCCTACCCAAGCAGGCATCAGATACAGGAAGTAACGATGCTCACAATAAAAAAGCA 303
Qy        112 LeuAspLys---------LeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
Db        304 GTTTCCAGATCAGCTGAACAGCAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCA 363
Qy        129 AlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheVal 148
Db        364 CAAGACATG---------------------------------------------------- 372
Qy        149 ProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsn 168
Db        373 ---ATTTCTGCTGGTGGAGAGAGTGTTGCTGGTATCACTGCAATATCT---GGCAAGCCG 426
Qy        169 AlaAspArgSerThrLeuProLysSerValGlnGluGly------AsnAspSerLysCys 186
Db        427 GGTGACAAGAAAAAAGAAAAGAAATCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCG 486
Qy        187 AsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAsp 206
Db        487 GATAAACCATCGGGAAAGTCAGGC---ATGGATGCTGCTTTGGATGACTTAATAGATACT 543
Qy        207 LeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSer 226
Db        544 TTAGGAGGACCTGAAGAA--------------------ACTGAAGAAGAAAATACAACG 582
Qy        227 ValAspValGlyAlaLeuProGluVal------ProGlnIleThrTrpHisIleGluVal 244
Db        583 TAT---------ACTGGACCAGAAGTTTCAGATCCAATGAGTTCCACCTACATAGAGGAA 633
Qy        245 AsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeuLysArg--- 263
Db        634 TTGGGTAAAAGAGAAGTCACAATTCCTCCAAAATATAGGGAACTATTGGCTAAAAAGGAA 693
Qy        263 ------------------------------------------------------------ 263
Db        694 GGGATCACAGGGCCTCCTGCAGACTCTTCGAAACCCATAGGGCCAGATGATGCTATAGAC 753
Qy        264 ---------------------------------------------AsnGluAspGluAsn 268
Db        754 GCCTTGTCATCTGACTTCACCTGTGGGTCGCCTACAGCTGCTGGAAAGAAAACTGAAAAA 813
Qy        269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
Db        814 GAGGAATCTACAGAAGTTTTAAAAGCTCAGTCAGCAGGGACAGTCAGAAGTGCTGCTCCA 873
Qy        289 MetSerGlyLysGluArg----------------------AspGlnValAlaGluGln 300
Db        874 CCCCAAGAGAAGAAAAGAAAGGTGGAGAAGGATACAATGAGTGATCAAGCACTCGAGGCT 933
Qy        301 -CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAs 320
Db        934 CTGTCGGCTTCACTGGGCACCC-----------GGCAAGCAG--------------AA 966
Qy        320 nGluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLy 340
Db        967 CCTGAGCTCGACCTCCGCTCAATTAAGGAAGTCGATGAGGCAAAAGCTAAAG------AA 1020
Qy        340 sLys-----------------LeuMetLysLysGlnGlnHisSerLysLysArgThrAl 354
Db       1021 GAAAAACTAGAGAAGTGTGGTGAGGATGATGAAACAATCCCATCT-GAGTACAGATTAAA 1079
Qy        354 aGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLys---------LysVa 371
```

```
Db       1080 ACCAGCCACGGATAAAGATGGAAAACCACTATTGCCAGAGCCTGAAGAAAAACCCAAGCC 1139

Qy        371 lArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVa 391

Db       1140 TCGGAGTGAATCAGAACTCATTGATGAACTTTCAGAAGATTTTGACCGGTCTGAA----- 1194

Qy        391 lHisArgGluAsnAlaAlaAspPro---------CysGluAspAspArgSerThrIlePr 408

Db       1195 -TGTAAAGAGAAACCATCTAAGCCAACTGAAAAGACAGAAGAATCTAAGGCCGCTGCTCC 1253

Qy        408 oValProMetGluValSerMet----------------------------------- 415

Db       1254 AGCTCCTGTGTCGGAGGCTGTGTGTCGGACCTCCATGTGTAGTATACAGTCAGCACCCCC 1313

Qy        416 -AspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLy 435

Db       1314 TGAGCCGGCTACCTTGAAGGGCACAGTGCCAGATGATGCTGTAGAAGCCTTGGCTGATAG 1373

Qy        435 sThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAs 455

Db       1374 CCTGGGGAAAAAGGAAGCAGAT---CCAGAAGATGGAAAACCTGTGATGGATAAAGTCAA 1430

Qy        455 nGlyLysLysLys--------------ArgThrGlySerValHisHisThrVal----- 468

Db       1431 GGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCTTGGTGAAAAAGAAGAAACAATTCCTCC 1490

Qy        469 ----------AlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSe 485

Db       1491 TGATTATAGATTAGAAGAGGTCAAGGATAAAGATGGAAAGCCACTCCTGCCAAAAGAGTC 1550

Qy        485 rThrGlnHis-----------------------------------------AspAspGl 491

Db       1551 TAAGGAACAGCTTCCACCCATGAGTGAAGACTTCCTTCTGGATGCTTTGTCTGAGGACTT 1610

Qy        491 uAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVa 511

Db       1611 CTCTGGTCCACAAAATGCTTCATCTCTTAAATTTGAAGATGCTAAACTTGCTGCTGCCAT 1670

Qy        511 lSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGl 531

Db       1671 CTCTGAAGTGGTTTCCCAA---------------------------------------- 1689

Qy        531 yLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAs 551

Db       1690 ----------ACCCCAGCTTCAACGACCCAAGCTGGAGCCCCACCCCGTGAT-------- 1731

Qy        551 nIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSe 571

Db       1732 ------------ACCTCGCAGAGTGACAAAGACCTCGATGATGCCTTGGATAAACTCTC 1778

Qy        571 rHisSer---AlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHi 590

Db       1779 TGACAGTCTAGGACAAAGGCAGCCTGACCCAGATGAGAACAAACCAATGGAAGAT----- 1833

Qy        590 sGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGl 610

Db       1834 ----------------AAAGTAAAGGAAAAAGCTAAAGCTGAA---CATAGAGACAAGCT 1874

Qy        610 nThrMetIleAspAsp---IleProMetAspIleValGluLeuLeuAlaLysAsnGlnHi 629

Db       1875 TGGAGAAAGAGATGACACTATCCCACCTGAATACAGACATCTCCTGGATGATAATGGACA 1934

Qy        629 sGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLysTh 649

Db       1935 GGACAAACCAGTGAAGCCACCTACAAAGAAATCAGAG-----------GATTCAAAGAA 1982

Qy        649 rThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSe 669

Db       1983 ACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCAGGAGATCTGGACAGCTGTCCCTC 2042

Qy        669 rValPheAspThrAsnSerGlnGln--------------LysSerLeuAlaSerGlnSe 684

Db       2043 CACTACAGAAACCTCACAGAACACAGCAAAGGATAAGTGCAAGAAGGCTGCTTCCAGCTC 2102

Qy        684 rThrGlnLysGluLeuGlnGlyHisLeu---------AlaLeuThrThrGlnGluSerPr 701

Db       2103 CAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAGCAAAGACAACAGAGGAAACTTC 2162

Qy        701 oHisProGlnAsn 705

Db       2163 CAAGCCAAAAGAT 2175
```

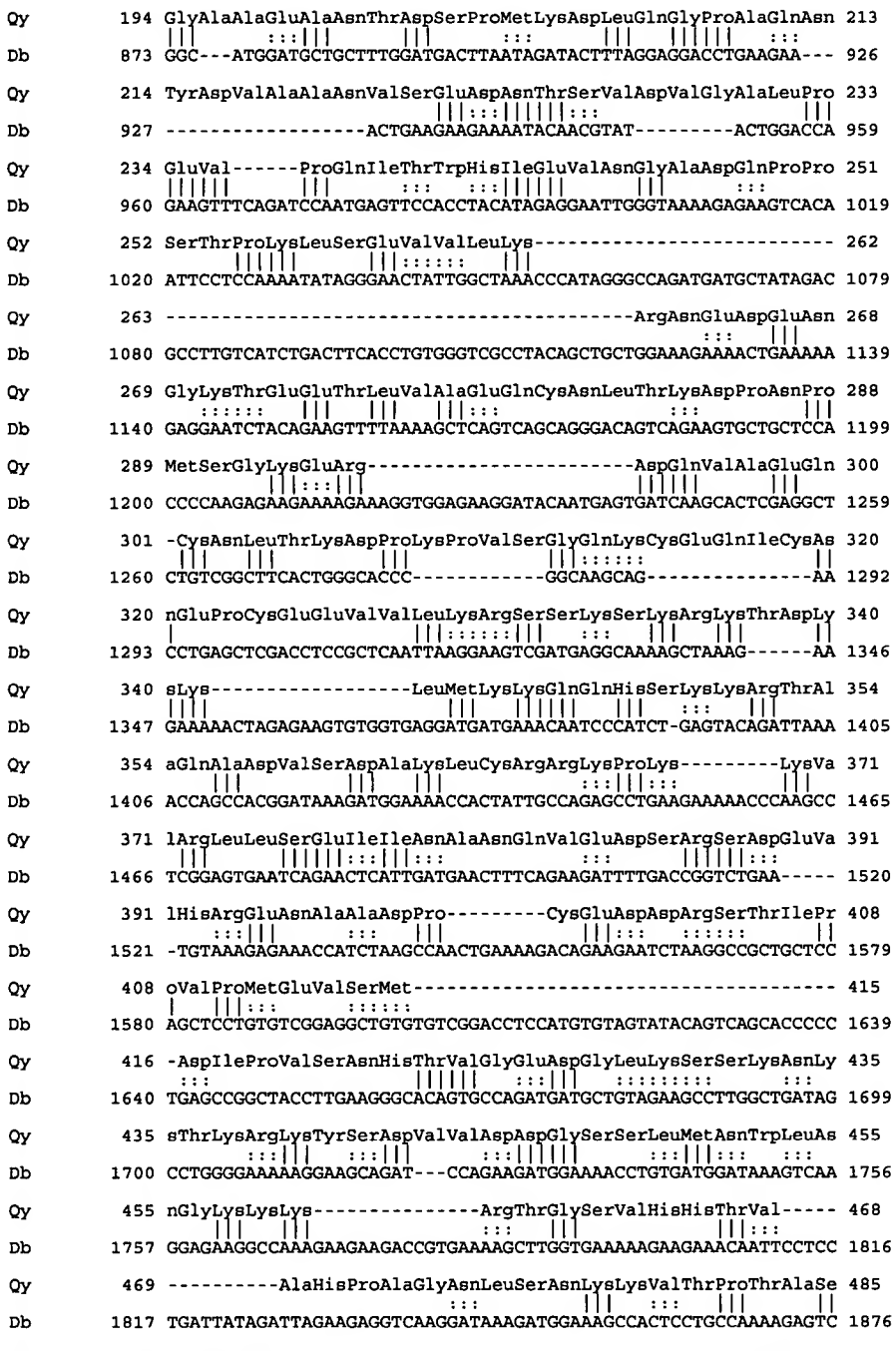
RESULT 10
US-11-124-368A-71
; Sequence 71, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 3076
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-71

Alignment Scores:

Pred. No.:	2.45e-06	Length:	3076
Score:	198.50	Matches:	163
Percent Similarity:	35.8%	Conservative:	104
Best Local Similarity:	21.9%	Mismatches:	288
Query Match:	3.6%	Indels:	191
DB:	12	Gaps:	34

US-09-828-068-2 (1-1057) x US-11-124-368A-71 (1-3076)

```
Qy         70 LysLysAspProLysPheCysSerLeuSer---------ArgIlePheHisAspGlnLys 86

Db        513 AAAACAGAACCTGAGAAGAAGTCACAGTCAACCAAGCTGTCTGTGGTTCATGAGAAAAAA 572

Qy         87 -----------LysCysAspGluHis--------------------LysAlaSerSer 95

Db        573 TCCCAAGAAGGAAAGCCAAAAGAACACACAGAGCCAAAAAGCCTACCCAAGCAGGCATCA 632

Qy         96 SerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeu 115

Db        633 GATACAGGAAGTAACGATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAG--- 689

Qy        116 LysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSer 135

Db        690 CAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCACAAGACATG----------- 737

Qy        136 AspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGln 155

Db        738 ----------------------------------------ATTTCTGCTGGTGGAGAG 755

Qy        156 LysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuPro 175

Db        756 AGTGTTGCTGGTATCACTGCAATATCT---GGCAAGCCGGGTGACAAGAAAAAAGAAAAG 812

Qy        176 LysSerValGlnGluGly------AsnAspSerLysCysAsnAlaProSerGlyLysAsn 193

Db        813 AAATCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCGGATAAACCATCGGGAAAGTCA 872
```



RESULT 11
US-11-124-368A-69
; Sequence 69, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 2943
; TYPE: DNA

; ORGANISM: Homo sapiens
US-11-124-368A-69

Alignment Scores:

Pred. No.:	3.1e-06	Length:	2943
Score:	197.00	Matches:	161
Percent Similarity:	35.8%	Conservative:	105
Best Local Similarity:	21.6%	Mismatches:	289
Query Match:	3.6%	Indels:	190
DB:	12	Gaps:	33

US-09-828-068-2 (1-1057) x US-11-124-368A-69 (1-2943)

```
Qy         70 LysLysAspProLysPheCysSerLeuSer---------ArgIlePheHisAspGlnLys 86
Db        536 AAAACAGAACCTGAGAAGAAGTCACAGTCAACCAAGCTGTCTGTGGTTCATGAGAAAAAA 595

Qy         87 ------------LysCysAspGluHis--------------------LysAlaSerSer 95
Db        596 TCCCAAGAAGGAAAGCCAAAAGAACACACAGAGCCAAAAAGCCTACCCAAGCAGGCATCA 655

Qy         96 SerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeu 115
Db        656 GATACAGGAAGTAACGATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAG--- 712

Qy        116 LysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSer 135
Db        713 CAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCACAAGACATG------------ 760

Qy        136 AspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGln 155
Db        761 ----------------------------------------ATTTCTGCTGGTGGAGAG 778

Qy        156 LysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuPro 175
Db        779 AGTGTTGCTGGTATCACTGCAATATCT---GGCAAGCCGGGTGACAAGAAAAAAGAAAAG 835

Qy        176 LysSerValGlnGluGly------AsnAspSerLysCysAsnAlaProSerGlyLysAsn 193
Db        836 AAATCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCGGATAAACCATCGGGAAAGTCA 895

Qy        194 GlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsn 213
Db        896 GGC---ATGGATGCTGCTTTGGATGACTTAATAGATACTTTAGGAGGACCTGAAGAA--- 949

Qy        214 TyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuPro 233
Db        950 -----------------ACTGAAGAAGAAAATACAACGTAT---------ACTGGACCA 982

Qy        234 GluVal------ProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
Db        983 GAAGTTTCAGATCCAATGAGTTCCACCTACATAGAGGAATTGGGTAAAAGAGAAGTCACA 1042

Qy        252 SerThrProLysLeuSerGluValValLeuLys-------------------------- 262
Db       1043 ATTCCTCCAAAATATAGGGAACTATTGGCTAAACCCATAGGGCCAGATGATGCTATAGAC 1102

Qy        263 ----------------------------------------ArgAsnGluAspGluAsn 268
Db       1103 GCCTTGTCATCTGACTTCACCTGTGGGTCGCCTACAGCTGCTGGAAAGAAAACTGAAAAA 1162

Qy        269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
Db       1163 GAGGAATCTACAGAAGTTTTAAAAGCTCAGTCAGCAGGGACAGTCAGAAGTGCTGCTCCA 1222

Qy        289 MetSerGlyLysGluArg-----------------------AspGlnValAlaGluGln 300
Db       1223 CCCCAAGAGAAGAAAAGAAAGGTGGAGAAGGATACAATGAGTGATCAAGCACTCGAGGCT 1282

Qy        301 -CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAs 320
Db       1283 CTGTCGGCTTCACTGGGCACCC-----------GGCAAGCAG--------------AA 1315
```

```
Qy        320 nGluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLy 340
Db       1316 CCTGAGCTCGACCTCCGCTCAATTAAGGAAGTCGATGAGGCAAAAGCTAAAG------AA 1369

Qy        340 sLys-----------------LeuMetLysLysGlnGlnHisSerLysLysArgThrAl 354
Db       1370 GAAAAACTAGAGAAGTGTGGTGAGGATGATGAAACAATCCCATCT-GAGTACAGATTAAA 1428

Qy        354 aGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLys---------LysVa 371
Db       1429 ACCAGCCACGGATAAAGATGGAAAACCACTATTGCCAGAGCCTGAAGAAAAACCCAAGCC 1488

Qy        371 lArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVa 391
Db       1489 TCGGAGTGAATCAGAACTCATTGATGAACTTTCAGAAGATTTTGACCGGTCTGAA----- 1543

Qy        391 lHisArgGluAsnAlaAlaAspPro---------CysGluAspAspArgSerThrIlePr 408
Db       1544 -TGTAAAGAGAAACCATCTAAGCCAACTGAAAAGACAGAAGAATCTAAGGCCGCTGCTCC 1602

Qy        408 oValProMetGluValSerMet------------------------------------ 415
Db       1603 AGCTCCTGTGTCGGAGGCTGTGTGTCGGACCTCCATGTGTAGTATACAGTCAGCACCCCC 1662

Qy        416 -AspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLy 435
Db       1663 TGAGCCGGCTACCTTGAAGGGCACAGTGCCAGATGATGCTGTAGAAGCCTTGGCTGATAG 1722

Qy        435 sThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAs 455
Db       1723 CCTGGGGAAAAAGGAAGCAGAT---CCAGAAGATGGAAAACCTGTGATGGATAAAGTCAA 1779

Qy        455 nGlyLysLysLys---------------ArgThrGlySerValHisHisThrVal----- 468
Db       1780 GGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCTTGGTGAAAAAGAAGAAACAATTCCTCC 1839

Qy        469 ----------AlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSe 485
Db       1840 TGATTATAGATTAGAAGAGGTCAAGGATAAAGATGGAAAGCCACTCCTGCCAAAAGAGTC 1899

Qy        485 rThrGlnHis-----------------------------------------AspAspGl 491
Db       1900 TAAGGAACAGCTTCCACCCATGAGTGAAGACTTCCTTCTGGATGCTTTGTCTGAGGACTT 1959

Qy        491 uAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVa 511
Db       1960 CTCTGGTCCACAAAATGCTTCATCTCTTAAATTTGAAGATGCTAAACTTGCTGCTGCCAT 2019

Qy        511 lSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGl 531
Db       2020 CTCTGAAGTGGTTTCCCAA---------------------------------------- 2038

Qy        531 yLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAs 551
Db       2039 ----------ACCCCAGCTTCAACGACCCAAGCTGGAGCCCCACCCCGTGAT-------- 2080

Qy        551 nIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSe 571
Db       2081 ------------ACCTCGAGTGACAAAGACCTCGATGATGCCTTGGATAAACTCTCTGA 2127

Qy        571 rHisSerAlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHisGl 591
Db       2128 CAGTCTAGGACAAAGGCAGCCTGACCCAGATGAGAACAAACCAATGGAAGAT-------- 2179

Qy        591 uGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGlnTh 611
Db       2180 ------------AAAGTAAAGGAAAAAGCTAAAGCTGAA---CATAGAGACAAGCTTGG 2223

Qy        611 rMetIleAspAsp---IleProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGl 630
Db       2224 AGAAAGAGATGACACTATCCCACCTGAATACAGACATCTCCTGGATGATAATGGACAGGA 2283

Qy        630 uArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLysThrTh 650
```

```
Db       2284 CAAACCAGTGAAGCCACCTACAAAGAAATCAGAG-----------GATTCAAAGAAACC 2331

Qy        650 rAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSerVa 670

Db       2332 TGCAGATGACCAAGACCCCATTGATGCTCTCTCAGGAGATCTGGACAGCTGTCCCTCCAC 2391

Qy        670 lPheAspThrAsnSerGlnGln--------------LysSerLeuAlaSerGlnSerTh 685

Db       2392 TACAGAAACCTCACAGAACACAGCAAAGGATAAGTGCAAGAAGGCTGCTTCCAGCTCCAA 2451

Qy        685 rGlnLysGluLeuGlnGlyHisLeu---------AlaLeuThrThrGlnGluSerProHi 702

Db       2452 AGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAGCAAAGACAACAGAGGAAACTTCCAA 2511

Qy        702 sProGlnAsn 705

Db       2512 GCCAAAAGAT 2521

RESULT 12
US-11-124-368A-76
; Sequence 76, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
;   LENGTH: 3122
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-76

Alignment Scores:
Pred. No.:              4.02e-06      Length:          3122
Score:                  196.00        Matches:         163
Percent Similarity:     35.4%         Conservative:    105
Best Local Similarity:  21.5%         Mismatches:      287
Query Match:            3.5%          Indels:          204
DB:                     12            Gaps:            34

US-09-828-068-2 (1-1057) x US-11-124-368A-76 (1-3122)

Qy         70 LysLysAspProLysPheCysSerLeuSer---------ArgIlePheHisAspGlnLys 86

Db        513 AAAACAGAACCTGAGAAGAAGTCACAGTCAACCAAGCTGTCTGTGGTTCATGAGAAAAAA 572

Qy         87 ------------LysCysAspGluHis---------------------LysAlaSerSer 95

Db        573 TCCCAAGAAGGAAAGCCAAAAGAACACACAGAGCCAAAAAGCCTACCCAAGCAGGCATCA 632

Qy         96 SerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeu 115

Db        633 GATACAGGAAGTAACGATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAG--- 689

Qy        116 LysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSer 135

Db        690 CAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCACAAGACATG------------ 737

Qy        136 AspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGln 155

Db        738 ----------------------------------------ATTTCTGCTGGTGGAGAG 755

Qy        156 LysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuPro 175

Db        756 AGTGTTGCTGGTATCACTGCAATATCT---GGCAAGCCGGGTGACAAGAAAAAAGAAAAG 812

Qy        176 LysSerValGlnGluGly------AsnAspSerLysCysAsnAlaProSerGlyLysAsn 193

Db        813 AAATCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCGGATAAACCATCGGGAAAGTCA 872

Qy        194 GlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsn 213

Db        873 GGC---ATGGATGCTGCTTTGGATGACTTAATAGATACTTTAGGAGGACCTGAAGAA--- 926

Qy        214 TyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuPro 233

Db        927 ------------------ACTGAAGAAGAAAATACAACGTAT---------ACTGGACCA 959

Qy        234 GluVal------ProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251

Db        960 GAAGTTTCAGATCCAATGAGTTCCACCTACATAGAGGAATTGGGTAAAAGAGAAGTCACA 1019

Qy        252 SerThrProLysLeuSerGluValValLeuLysArg------------------------ 263

Db       1020 ATTCCTCCAAAATATAGGGAACTATTGGCTAAAAAGGAAGGGATCACAGGGCCTCCTGCA 1079

Qy        263 ------------------------------------------------------------ 263

Db       1080 GACTCTTCGAAACCCATAGGGCCAGATGATGCTATAGACGCCTTGTCATCTGACTTCACC 1139

Qy        264 ----------------------AsnGluAspGluAsnGlyLysThrGluGluThrLeu 275

Db       1140 TGTGGGTCGCCTACAGCTGCTGGAAAGAAAACTGAAAAAGAGGAATCTACAGAAGTTTTA 1199

Qy        276 ValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGlyLysGluArg--- 294

Db       1200 AAAGCTCAGTCAGCAGGGACAGTCAGAAGTGCTGCTCCACCCCAAGAGAAGAAAAGAAAG 1259

Qy        295 ---------------------AspGlnValAlaGluGln-CysAsnLeuThrLysAspPr 307

Db       1260 GTGGAGAAGGATACAATGAGTGATCAAGCACTCGAGGCTCTGTCGGCTTCACTGGGCACC 1319

Qy        307 oLysProValSerGlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValVa 327

Db       1320 C------------GGCAAGCAG--------------AACCTGAGCTCGACCTCCGCTCA 1352

Qy        327 lLeuLysArgSerSerLysSerLysArgLysThrAspLysLys----------------- 341

Db       1353 ATTAAGGAAGTCGATGAGGCAAAAGCTAAAG------AAGAAAAACTAGAGAAGTGTGGT 1406

Qy        342 -LeuMetLysLysGlnGlnHisSerLysLysArgThrAlaGlnAlaAspValSerAspAl 361

Db       1407 GAGGATGATGAAACAATCCCATCT-GAGTACAGATTAAAACCAGCCACGGATAAAGATGG 1465

Qy        361 aLysLeuCysArgArgLysProLys---------LysValArgLeuLeuSerGluIleIl 378

Db       1466 AAAACCACTATTGCCAGAGCCTGAAGAAAAACCCAAGCCTCGGAGTGAATCAGAACTCAT 1525

Qy        378 eAsnAlaAsnGlnValGluAspSerArgSerAspGluValHisArgGluAsnAlaAlaAs 398

Db       1526 TGATGAACTTTCAGAAGATTTTGACCGGTCTGAA------TGTAAAGAGAAACCATCTAA 1579

Qy        398 pPro---------CysGluAspAspArgSerThrIleProValProMetGluValSerMe 415

Db       1580 GCCAACTGAAAAGACAGAAGAATCTAAGGCCGCTGCTCCAGCTCCTGTGTCGGAGGCTGT 1639

Qy        415 t-------------------------------------AspIleProValSerAsnHi 422

Db       1640 GTGTCGGACCTCCATGTGTAGTATACAGTCAGCACCCCCTGAGCCGGCTACCTTGAAGGG 1699

Qy        422 sThrValGlyGluAspGlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAs 442

Db       1700 CACAGTGCCAGATGATGCTGTAGAAGCCTTGGCTGATAGCCTGGGGAAAAAGGAAGCAGA 1759
```

```
Qy        442 pValValAspAspGlySerSerLeuMetAsnTrpLeuAsnGlyLysLysLys-------- 459
Db       1760 T---CCAGAAGATGGAAAACCTGTGATGGATAAAGTCAAGGAGAAGGCCAAAGAAGAAGA 1816

Qy        460 -------ArgThrGlySerValHisHisThrVal--------------AlaHisProAl 472
Db       1817 CCGTGAAAAGCTTGGTGAAAAAGAAGAAACAATTCCTCCTGATTATAGATTAGAAGAGGT 1876

Qy        472 aGlyAsnLeuSerAsnLysLysValThrProThrAlaSerThrGlnHis---------- 488
Db       1877 CAAGGATAAAGATGGAAAGCCACTCCTGCCAAAAGAGTCTAAGGAACAGCTTCCACCCAT 1936

Qy        489 ------------------------------AspAspGluAsnAspThrGluAsnGlyLe 498
Db       1937 GAGTGAAGACTTCCTTCTGGATGCTTTGTCTGAGGACTTCTCTGGTCCACAAAATGCTTC 1996

Qy        498 uAspThrAsnMetHisLysThrAspValCysGlnHisValSerGluIleSerThrGlnAr 518
Db       1997 ATCTCTTAAATTTGAAGATGCTAAACTTGCTGCTGCCATCTCTGAAGTGGTTTCCCAA-- 2054

Qy        518 gCysSerSerLysGlyLysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSe 538
Db       2055 ------------------------------------------------ACCCCAGCTTC 2065

Qy        538 rThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGl 558
Db       2066 AACGACCCAAGCTGGAGCCCCACCCCGTGAT--------------------ACCTCGCA 2104

Qy        558 uAspGlnCysGlnMetGluThrGluAsnSerValLeuSerHisSer---AlaLysValSe 577
Db       2105 GAGTGACAAAGACCTCGATGATGCCTTGGATAAACTCTCTGACAGTCTAGGACAAAGGCA 2164

Qy        577 rProAlaGluHisAspIleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLy 597
Db       2165 GCCTGACCCAGATGAGAACAAACCAATGGAAGAT--------------------AAAGT 2203

Qy        597 sLysLysLysGlnLysLeuGluValThrArgGluLysGlnThrMetIleAspAsp---Il 616
Db       2204 AAAGGAAAAAGCTAAAGCTGAA---CATAGAGACAAGCTTGGAGAAAGAGATGACACTAT 2260

Qy        616 eProMetAspIleValGluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGl 636
Db       2261 CCCACCTGAATACAGACATCTCCTGGATGATAATGGACAGGACAAACCAGTGAAGCCACC 2320

Qy        636 uThrAspCysSerAspIleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysVa 656
Db       2321 TACAAAGAAATCAGAG------------GATTCAAAGAAACCTGCAGATGACCAAGACCC 2368

Qy        656 lIleValAlaAlaLysAspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGl 676
Db       2369 CATTGATGCTCTCTCAGGAGATCTGGACAGCTGTCCCTCCACTACAGAAACCTCACAGAA 2428

Qy        676 nGln--------------LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGl 691
Db       2429 CACAGCAAAGGATAAGTGCAAGAAGGCTGCTTCCAGCTCCAAAGCACCTAAGAATGGAGG 2488

Qy        691 yHisLeu---------AlaLeuThrThrGlnGluSerProHisProGlnAsn 705
Db       2489 TAAAGCGAAGGATTCAGCAAAGACAACAGAGGAAACTTCCAAGCCAAAAGAT 2540
```

RESULT 13
US-10-821-234-625
; Sequence 625, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 625
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-625

Alignment Scores:

Pred. No.:	2.96e-06	Length:	2004
Score:	195.00	Matches:	159
Percent Similarity:	34.5%	Conservative:	98
Best Local Similarity:	21.3%	Mismatches:	285
Query Match:	3.5%	Indels:	204
DB:	8	Gaps:	32

US-09-828-068-2 (1-1057) x US-10-821-234-625 (1-2004)

```
Qy         69 GlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLysCys 88
Db         52 CAGTCAACCAAGCCAAAA------AGCCTACCCAAGCAGGCATCAGATACAGGAAGTAAC 105

Qy         89 AspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCys 108
Db        106 GATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAG----------------- 147

Qy        109 SerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128
Db        148 --------------------CAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCA 186

Qy        129 AlaLysGlnAsnGlyThrSerAspGlyCysSerIleThrPheValArgSerThrPheVal 148
Db        187 CAAGACATG---------------------------------------------------- 195

Qy        149 ProAlaSerValGlySerGlnLysValSerProSerThrGlnSerSerGlnGlyLysAsn 168
Db        196 ---ATTTCTGCTGGTGGAGAGAGTGTTGCTGGTATCACTGCAATATCT---GGCAAGCCG 249

Qy        169 AlaAspArgSerThrLeuProLysSerValGlnGluGly------AsnAspSerLysCys 186
Db        250 GGTGACAAGAAAAAAGAAAAGAAATCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCG 309

Qy        187 AsnAlaProSerGlyLysAsnGlyAlaAlaGluAlaAsnThrAspSerProMetLysAsp 206
Db        310 GATAAACCATCGGGAAAGTCAGGC---ATGGATGCTGCTTTGGATGACTTAATAGATACT 366

Qy        207 LeuGlnGlyProAlaGlnAsnTyrAspValAlaAlaAsnValSerGluAspAsnThrSer 226
Db        367 TTAGGAGGACCTGAAGAA--------------------ACTGAAGAAGAAAATACAACG 405

Qy        227 ValAspValGlyAlaLeuProGluVal------ProGlnIleThrTrpHisIleGluVal 244
Db        406 TAT---------ACTGGACCAGAAGTTTCAGATCCAATGAGTTCCACCTACATAGAGGAA 456

Qy        245 AsnGlyAlaAspGlnProProSerThrProLysLeuSerGluValValLeuLysArg--- 263
Db        457 TTGGGTAAAAGAGAAGTCACAATTCCTCCAAAATATAGGGAACTATTGGCTAAAAAGGAA 516

Qy        263 ------------------------------------------------------------ 263
Db        517 GGGATCACAGGGCCTCCTGCAGACTCTTCGAAACCCATAGGGCCAGATGATGCTATAGAC 576

Qy        264 ---------------------------------------------AsnGluAspGluAsn 268
Db        577 GCCTTGTCATCTGACTTCACCTGTGGGTCGCCTACAGCTGCTGGAAAGAAAACTGAAAAA 636

Qy        269 GlyLysThrGluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnPro 288
Db        637 GAGGAATCTACAGAAGTTTTAAAAGCTCAGTCAGCAGGGACAGTCAGAAGTGCTGCTCCA 696

Qy        289 MetSerGlyLysGluArg----------------------AspGlnValAlaGluGln 300
```

```
Db        697 CCCCAAGAGAAGAAAAGAAAGGTGGAGAAGGATACAATGAGTGATCAAGCACTCGAGGCT 756

Qy        301 -CysAsnLeuThrLysAspProLysProValSerGlyGlnLysCysGluGlnIleCysAs 320

Db        757 CTGTCGGCTTCACTGGGCACCC-----------GGCAAGCAG--------------AA 789

Qy        320 nGluProCysGluGluValValLeuLysArgSerSerLysSerLysArgLysThrAspLy 340

Db        790 CCTGAGCTCGACCTCCGCTCAATTAAGGAAGTCGATGAGGCAAAAGCTAAAG------AA 843

Qy        340 sLys-----------------LeuMetLysLysGlnGlnHisSerLysLysArgThrAl 354

Db        844 GAAAAACTAGAGAAGTGTGGTGAGGATGATGAAACAATCCCATCT-GAGTACAGATTAAA 902

Qy        354 aGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLys---------LysVa 371

Db        903 ACCAGCCACGGATAAAGATGGAAAACCACTATTGCCAGAGCCTGAAGAAAAACCCAAGCC 962

Qy        371 lArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGluAspSerArgSerAspGluVa 391

Db        963 TCGGAGTGAATCAGAACTCATTGATGAACTTTCAGAAGATTTTGACCGGTCTGAA----- 1017

Qy        391 lHisArgGluAsnAlaAlaAspPro---------CysGluAspAspArgSerThrIlePr 408

Db       1018 -TGTAAAGAGAAACCATCTAAGCCAACTGAAAAGACAGAAGAATCTAAGGCCGCTGCTCC 1076

Qy        408 oValProMetGluValSerMet------------------------------------ 415

Db       1077 AGCTCCTGTGTCGGAGGCTGTGTGTCGGACCTCCATGTGTAGTATACAGTCAGCACCCCC 1136

Qy        416 -AspIleProValSerAsnHisThrValGlyGluAspGlyLeuLysSerSerLysAsnLy 435

Db       1137 TGAGCCGGCTACCTTGAAGGGCACAGTGCCAGATGATGCTGTAGAAGCCTTGGCTGATAG 1196

Qy        435 sThrLysArgLysTyrSerAspValValAspAspGlySerSerLeuMetAsnTrpLeuAs 455

Db       1197 CCTGGGGAAAAAGGAAGCAGAT---CCAGAAGATGGAAAACCTGTGATGGATAAAGTCAA 1253

Qy        455 nGlyLysLysLys--------------ArgThrGlySerValHisHisThrVal----- 468

Db       1254 GGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCTTGGTGAAAAAGAAGAAACAATTCCTCC 1313

Qy        469 ----------AlaHisProAlaGlyAsnLeuSerAsnLysLysValThrProThrAlaSe 485

Db       1314 TGATTATAGATTAGAAGAGGTCAAGGATAAAGATGGAAAGCCACTCCTGCCAAAAGAGTC 1373

Qy        485 rThrGlnHis-------------------------------------------AspAspGl 491

Db       1374 TAAGGAACAGCTTCCACCCATGAGTGAAGACTTCCTTCTGGATGCTTTGTCTGAGGACTT 1433

Qy        491 uAsnAspThrGluAsnGlyLeuAspThrAsnMetHisLysThrAspValCysGlnHisVa 511

Db       1434 CTCTGGTCCACAAAATGCTTCATCTCTTAAATTTGAAGATGCTAAACTTGCTGCTGCCAT 1493

Qy        511 lSerGluIleSerThrGlnArgCysSerSerLysGlyLysThrAlaGlyLeuSerLysGl 531

Db       1494 CTCTGAAGTGGTTTCCCAA---------------------------------------- 1512

Qy        531 yLysThrHisSerAlaAlaSerThrLysTyrGlyGlyGluSerThrArgAsnGlyGlnAs 551

Db       1513 ----------ACCCCAGCTTCAACGACCCAAGCTGGAGCCCCACCCCGTGAT-------- 1554

Qy        551 nIleHisValLeuSerAlaGluAspGlnCysGlnMetGluThrGluAsnSerValLeuSe 571

Db       1555 ------------ACCTCGCAGAGTGACAAAGACCTCGATGATGCCTTGGATAAACTCTC 1601

Qy        571 rHisSer---AlaLysValSerProAlaGluHisAspIleGlnIleMetSerAspLeuHi 590

Db       1602 TGACAGTCTAGGACAAAGGCAGCCTGACCCAGATGAGAACAAACCAATGGAAGAT----- 1656

Qy        590 sGluGlnSerLeuProLysLysLysLysGlnLysLeuGluValThrArgGluLysGl 610

Db       1657 ---------------AAAGTAAAGGAAAAAGCTAAAGCTGAA---CATAGAGACAAGCT 1697

Qy        610 nThrMetIleAspAsp---IleProMetAspIleValGluLeuLeuAlaLysAsnGlnHi 629

Db       1698 TGGAGAAAGAGATGACACTATCCCACCTGAATACAGACATCTCCTGGATGATAATGGACA 1757

Qy        629 sGluArgGlnLeuMetThrGluThrAspCysSerAspIleAsnArgIleGlnSerLysTh 649

Db       1758 GGACAAACCAGTGAAGCCACCTACAAAGAAATCAGAG-----------GATTCAAAGAA 1805

Qy        649 rThrAlaAspAspAspCysValIleValAlaAlaLysAspGlySerAspTyrAlaSerSe 669

Db       1806 ACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCAGGAGATCTGGACAGCTGTCCCTC 1865

Qy        669 rValPheAspThrAsnSerGlnGln--------------LysSerLeuAlaSerGlnSe 684

Db       1866 CACTACAGAAACCTCACAGAACACAGCAAAGGATAAGTGCAAGAAGGCTGCTTCCAGCTC 1925

Qy        684 rThrGlnLysGluLeuGlnGlyHisLeu---------AlaLeuThrThrGlnGluSerPr 701

Db       1926 CAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCAGCAAAGACAACAGAGGAAACTTC 1985

Qy        701 oHisProGlnAsn 705

Db       1986 CAAGCCAAAAGAT 1998
```

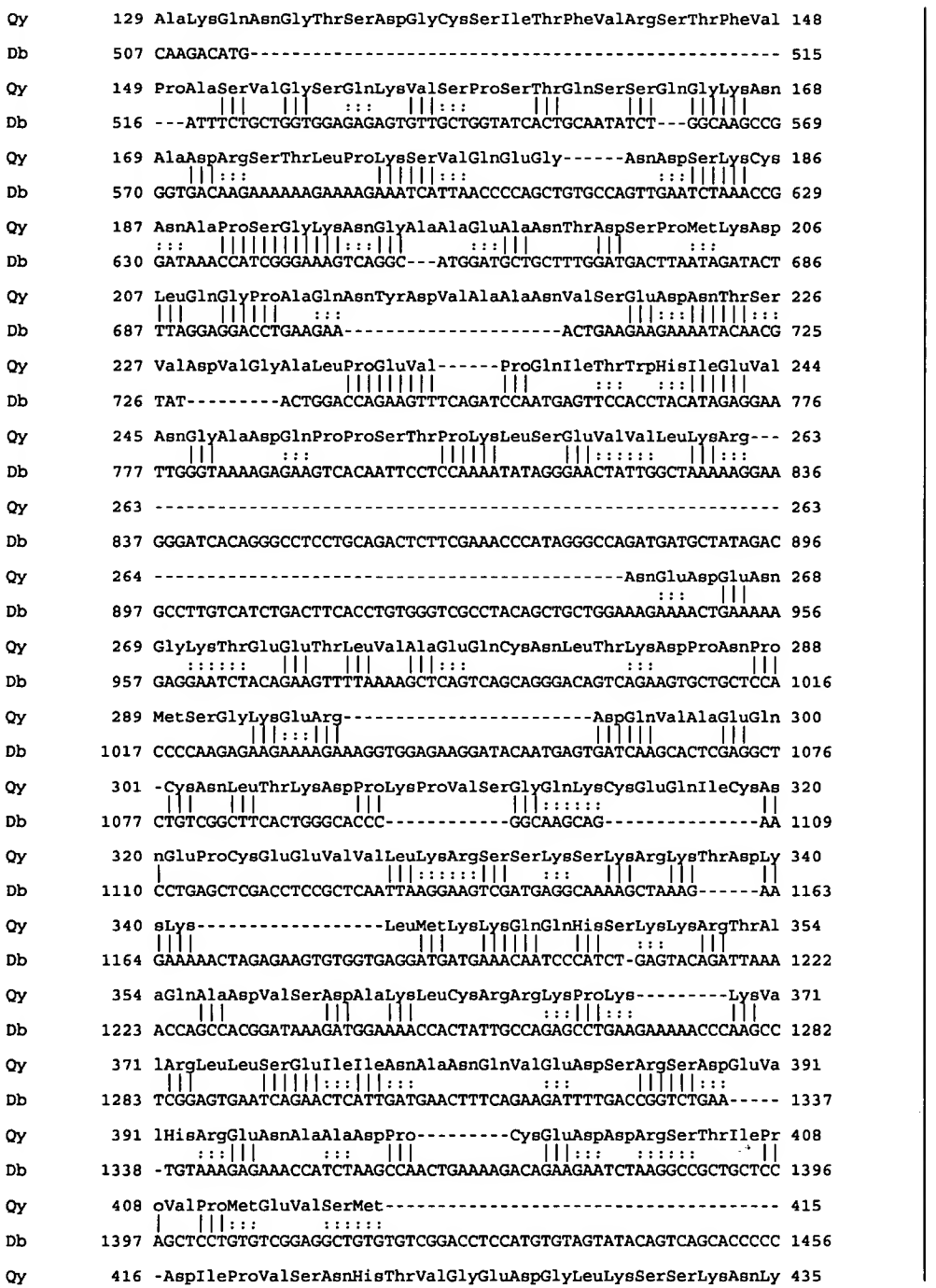
RESULT 14
US-11-124-368A-80
; Sequence 80, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 4075
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-80

Alignment Scores:

Pred. No.:	6.55e-06	Length:	4075
Score:	195.00	Matches:	159
Percent Similarity:	34.5%	Conservative:	98
Best Local Similarity:	21.3%	Mismatches:	285
Query Match:	3.5%	Indels:	204
DB:	12	Gaps:	32

US-09-828-068-2 (1-1057) x US-11-124-368A-80 (1-4075)

```
Qy         69 GlnLysLysAspProLysPheCysSerLeuSerArgIlePheHisAspGlnLysLysCys 88

Db        372 CAGTCAACCAAGCCAAAA------AGCCTACCCAAGCAGGCATCAGATACAGGAAGTAAC 425

Qy         89 AspGluHisLysAlaSerSerSerProPheSerValAlaLysPheArgArgTrpAspCys 108

Db        426 GATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAG------------------ 467

Qy        109 SerLysCysLeuAspLysLeuLysThrSerAspAsnGlyThrAlaProArgThrLeuPro 128

Db        468 --------------------CAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCA 506
```



RESULT 15
US-11-124-368A-74
; Sequence 74, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill

```
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
;   LENGTH: 2791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-74

Alignment Scores:
Pred. No.:              1.35e-05        Length:         2791
Score:                  189.00          Matches:        149
Percent Similarity:     35.7%           Conservative:   98
Best Local Similarity:  21.5%           Mismatches:     265
Query Match:            3.4%            Indels:         180
DB:                     12              Gaps:           33

US-09-828-068-2 (1-1057) x US-11-124-368A-74 (1-2791)

Qy          70 LysLysAspProLysPheCysSerLeuSer---------ArgIlePheHisAspGlnLys 86
               |||   :::|||:::       |||   |||             :::   |||:::::|||
Db         513 AAAACAGAACCTGAGAAGAAGTCACAGTCAACCAAGCTGTCTGTGGTTCATGAGAAAAAA 572

Qy          87 ------------LysCysAspGluHis--------------------LysAlaSerSer 95
                         |||       ||||||                   |||   :::|||
Db         573 TCCCAAGAAGGAAAGCCAAAAGAACACACAGAGCCAAAAAGCCTACCCAAGCAGGCATCA 632

Qy          96 SerProPheSerValAlaLysPheArgArgTrpAspCysSerLysCysLeuAspLysLeu 115
                  |||             :::          |||:::      ::::::
Db         633 GATACAGGAAGTAACGATGCTCACAATAAAAAAGCAGTTTCCAGATCAGCTGAACAG--- 689

Qy         116 LysThrSerAspAsnGlyThrAlaProArgThrLeuProAlaLysGlnAsnGlyThrSer 135
               :::   |||:::       |||   |||:::|||   |||
Db         690 CAGCCATCAGAGAAATCAACAGAACCAAAGACTAAACCACAAGACATG------------ 737

Qy         136 AspGlyCysSerIleThrPheValArgSerThrPheValProAlaSerValGlySerGln 155
                                                        |||   |||   :::
Db         738 ----------------------------------------ATTTCTGCTGGTGGAGAG 755

Qy         156 LysValSerProSerThrGlnSerSerGlnGlyLysAsnAlaAspArgSerThrLeuPro 175
                  |||:::      |||      |||   ||||||      |||:::
Db         756 AGTGTTGCTGGTATCACTGCAATATCT---GGCAAGCCGGGTGACAAGAAAAAAGAAAAG 812

Qy         176 LysSerValGlnGluGly------AsnAspSerLysCysAsnAlaProSerGlyLysAsn 193
               ||||||:::                    :::|||||||   :::   |||||||||||:::
Db         813 AAATCATTAACCCCAGCTGTGCCAGTTGAATCTAAACCGGATAAACCATCGGGAAAGTCA 872

Qy         194 GlyAlaAlaGluAlaAsnThrAspSerProMetLysAspLeuGlnGlyProAlaGlnAsn 213
               |||      :::|||      |||      :::      |||   ||||||   :::
Db         873 GGC---ATGGATGCTGCTTTGGATGACTTAATAGATACTTTAGGAGGACCTGAAGAA--- 926

Qy         214 TyrAspValAlaAlaAsnValSerGluAspAsnThrSerValAspValGlyAlaLeuPro 233
                                    |||:::||||||:::                  |||
Db         927 -----------------ACTGAAGAAGAAAATACAACGTAT---------ACTGGACCA 959

Qy         234 GluVal------ProGlnIleThrTrpHisIleGluValAsnGlyAlaAspGlnProPro 251
               ||||||      |||      :::   :::||||||      |||      :::
Db         960 GAAGTTTCAGATCCAATGAGTTCCACCTACATAGAGGAATTGGGTAAAAGAGAAGTCACA 1019

Qy         252 SerThrProLysLeuSerGluValValLeuLysArgAsnGluAspGluAsnGlyLysThr 271
                  ||||||      |||::::::   |||:::
Db        1020 ATTCCTCCAAAATATAGGGAACTATTGGCTAAAAAG---------------------- 1055

Qy         272 GluGluThrLeuValAlaGluGlnCysAsnLeuThrLysAspProAsnProMetSerGly 291
                  |||   :::                 :::   :::|||          |||
Db        1056 ---GAAGGGATCACAGGGCCTCCTGCAGACTCTTCGAAACCCATAGGGCCAGATGATGCT 1112

Qy         292 LysGluArgAspGlnValAlaGluGlnCysAsnLeuThrLysAspProLysProValSer 311
                  :::             |||          :::   |||                  :::
Db        1113 ATAGAC------------GCCTTGTCATCTGACTTCACCTGTGGGTCGCCTACAGCTGCT 1160

Qy         312 GlyGlnLysCysGluGlnIleCysAsnGluProCysGluGluValValLeuLysArgSer 331
               |||:::|||   |||:::         |||        ||||||:::      :::|||
Db        1161 GGAAAGAAAACTGAAAAA---------GAGGAATCTACAGAAGTTTTAAAAGCTCAGTCA 1211

Qy         332 SerLysSerLysArgLysThrAspLysLysLeuMetLysLysGlnGlnHisSerLysLys 351
               :::   :::   |||                                  |||      |||:::
Db        1212 GCAGGGACAGTCAGAAGTGCT--------------GCTCCACCCCAAGAGAAGAAAAGA 1256

Qy         352 ArgThrAlaGlnAlaAspValSerAspAlaLysLeuCysArgArgLysProLysLysVal 371
               :::      :::      :::||||||      |||
Db        1257 AAGGTGGAGAAGGATACAATGAGTGATCAAGCACTC---------------------- 1292

Qy         372 ArgLeuLeuSerGluIleIleAsnAlaAsnGlnValGlu---------AspSerArgSer 388
                  ||||||      :::        |||   |||          |||   ||||||
Db        1293 GAGGCTCTGTCGGCTTCACTGGGCACCCGGCAAGCAGAACCTGAGCTCGACCTCCGCTCA 1352

Qy         389 ---AspGluValHisArgGluAsnAlaAlaAspProCysGluAspAspArgSerThrIle 407
                  ||||||                        |||::::::   |||:::
Db        1353 ATTAAGGAAGTC----------------------------GATGAGAAGGGCACAGTG 1382

Qy         408 ProValProMetGluValSerMetAspIleProValSerAsnHisThrValGlyGluAsp 427
               |||                                                  :::|||
Db        1383 CCA-------------------------------------------------GATGAT 1391

Qy         428 GlyLeuLysSerSerLysAsnLysThrLysArgLysTyrSerAspValValAspAspGly 447
                  :::::::::      :::        :::|||   :::|||      :::||||||
Db        1392 GCTGTAGAAGCCTTGGCTGATAGCCTGGGGAAAAAGGAAGCAGAT---CCAGAAGATGGA 1448

Qy         448 SerSerLeuMetAsnTrpLeuAsnGlyLysLysLys---------------ArgThrGly 462
                  :::|||:::   :::       |||   |||                :::   |||
Db        1449 AAACCTGTGATGGATAAAGTCAAGGAGAAGGCCAAAGAAGAAGACCGTGAAAAGCTTGGT 1508

Qy         463 SerValHisHisThrVal---------------AlaHisProAlaGlyAsnLeuSerAsn 477
                        |||:::                        :::
Db        1509 GAAAAAGAAGAAACAATTCCTCCTGATTATAGATTAGAAGAGGTCAAGGATAAAGATGGA 1568

Qy         478 LysLysValThrProThrAlaSerThrGlnHis---------------------------- 488
               |||   :::   |||       |||   :::
Db        1569 AAGCCACTCCTGCCAAAAGAGTCTAAGGAACAGCTTCCACCCATGAGTGAAGACTTCCTT 1628

Qy         489 --------------AspAspGluAsnAspThrGluAsnGlyLeuAspThrAsnMetHis 503
                           :::|||   :::      ::::||
Db        1629 CTGGATGCTTTGTCTGAGGACTTCTCTGGTCCACAAAATGCTTCATCTCTTAAATTTGAA 1688

Qy         504 LysThrAspValCysGlnHisValSerGluIleSerThrGlnArgCysSerSerLysGly 523
                  :::             :::||||||:::   :::|||
Db        1689 GATGCTAAACTTGCTGCTGCCATCTCTGAAGTGGTTTCCCAA----------------- 1730

Qy         524 LysThrAlaGlyLeuSerLysGlyLysThrHisSerAlaAlaSerThrLysTyrGlyGly 543
                                          :::   |||||||||          |||
Db        1731 -------------------------------ACCCCAGCTTCAACGACCCAAGCTGGA 1757

Qy         544 GluSerThrArgAsnGlyGlnAsnIleHisValLeuSerAlaGluAspGlnCysGlnMet 563
                        |||:::                    :::::::::      :::
Db        1758 GCCCCACCCCGTGAT--------------------ACCTCGCAGAGTGACAAAGACCTC 1796

Qy         564 GluThrGluAsnSerValLeuSerHisSer---AlaLysValSerProAlaGluHisAsp 582
               :::                |||||   |||      :::      |||          :::
Db        1797 GATGATGCCTTGGATAAACTCTCTGACAGTCTAGGACAAAGGCAGCCTGACCCAGATGAG 1856

Qy         583 IleGlnIleMetSerAspLeuHisGluGlnSerLeuProLysLysLysLysGlnLys 602
               :::   |||   |||                     |||   |||:::|||   |||
Db        1857 AACAAACCAATGGAAGAT--------------------AAAGTAAAGGAAAAAGCTAAA 1895

Qy         603 LeuGluValThrArgGluLysGlnThrMetIleAspAsp---IleProMetAspIleVal 621
```

```
Db       1896 GCTGAA---CATAGAGACAAGCTTGGAGAAAGAGATGACACTATCCCACCTGAATACAGA 1952

Qy        622 GluLeuLeuAlaLysAsnGlnHisGluArgGlnLeuMetThrGluThrAspCysSerAsp 641

Db       1953 CATCTCCTGGATGATAATGGACAGGACAAACCAGTGAAGCCACCTACAAAGAAATCAGAG 2012

Qy        642 IleAsnArgIleGlnSerLysThrThrAlaAspAspAspCysValIleValAlaAlaLys 661

Db       2013 ------------GATTCAAAGAAACCTGCAGATGACCAAGACCCCATTGATGCTCTCTCA 2060

Qy        662 AspGlySerAspTyrAlaSerSerValPheAspThrAsnSerGlnGln------------ 677

Db       2061 GGAGATCTGGACAGCTGTCCCTCCACTACAGAAACCTCACAGAACACAGCAAAGGATAAG 2120

Qy        678 ---LysSerLeuAlaSerGlnSerThrGlnLysGluLeuGlnGlyHisLeu--------- 693

Db       2121 TGCAAGAAGGCTGCTTCCAGCTCCAAAGCACCTAAGAATGGAGGTAAAGCGAAGGATTCA 2180

Qy        694 AlaLeuThrThrGlnGluSerProHisProGlnAsn 705

Db       2181 GCAAAGACAACAGAGGAAACTTCCAAGCCAAAAGAT 2216
```

Search completed: March 7, 2006, 14:44:25
Job time : 711 secs